

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:22:20 ; Search time 75 Seconds
(without alignments)
1188.596 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRGRSLSPSRDKLFPN.....RLADSRNPAGRFSTQEEIQ 669

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3470	100.0	669	21 AA71171	Human Heregulin HR
2	3470	100.0	669	22 AAU09886	Human Heregulin, H
3	3466	99.9	669	13 AAR29570	Human heregulin-al
4	3283.5	94.6	675	13 AAR29571	Human heregulin-be
5	3283.5	94.6	675	19 AA74434	Amino acid sequenc
6	3283.5	94.6	675	21 AA71172	Human Heregulin HR
7	3283.5	94.6	675	21 AA71199	Human Heregulin HR
8	3283.5	94.6	675	21 AA71200	Human Heregulin HR
9	3283.5	94.6	675	21 AA71203	Human Heregulin, H
10	3283.5	94.6	675	22 AAU09887	Human Heregulin, H

11	3282.5	94.6	675	21 AA71202	Human Heregulin HR
12	3281.5	94.6	675	19 AA74433	Amino acid sequenc
13	3280.5	94.5	675	21 AA71188	Human Heregulin HR
14	3280.5	94.5	675	21 AA71189	Human Heregulin HR
15	3280.5	94.5	675	21 AA71191	Human Heregulin HR
16	3280.5	94.5	675	21 AA71198	Human Heregulin HR
17	3279.5	94.5	675	21 AA71182	Human Heregulin HR
18	3279.5	94.5	675	21 AA71194	Human Heregulin HR
19	3278.5	94.5	675	21 AA71180	Human Heregulin HR
20	3278.5	94.5	675	21 AA71185	Human Heregulin HR
21	3278.5	94.5	675	21 AA71197	Human Heregulin HR
22	3277.5	94.5	675	21 AA71181	Human Heregulin HR
23	3277.5	94.5	675	21 AA71183	Human Heregulin HR
24	3277.5	94.5	675	21 AA71184	Human Heregulin HR
25	3277.5	94.5	675	21 AA71186	Human Heregulin HR
26	3277.5	94.5	675	21 AA71187	Human Heregulin HR
27	3277.5	94.5	675	21 AA71190	Human Heregulin HR
28	3277.5	94.5	675	21 AA71201	Human Heregulin HR
29	3276.5	94.4	675	21 AA71178	Human Heregulin HR
30	3276.5	94.4	675	21 AA71192	Human Heregulin HR
31	3276.5	94.4	675	21 AA71195	Human Heregulin HR
32	3275.5	94.4	675	21 AA71196	Human Heregulin HR
33	3274.5	94.4	675	21 AA71179	Human Heregulin HR
34	3274.5	94.4	675	21 AA71193	Human Heregulin HR
35	3268.5	94.2	675	19 AA74509	Amino acid sequenc
36	3267.5	94.0	675	19 AA74503	Amino acid sequenc
37	3262.5	94.0	675	19 AA74488	Amino acid sequenc
38	3259.5	93.9	675	19 AA74500	Amino acid sequenc
39	3252.5	93.7	675	19 AA74487	Amino acid sequenc
40	3252.5	93.7	675	19 AA74436	Amino acid sequenc
41	3252.5	93.7	675	19 AA74491	Amino acid sequenc
42	3252.5	93.7	675	19 AA74489	Amino acid sequenc
43	3250.5	93.7	675	19 AA74490	Amino acid sequenc
44	3249.5	93.6	675	19 AA74486	Amino acid sequenc
45	3247.5	93.6	675	19 AA74492	Amino acid sequenc

ALIGNMENTS

RESULT 1

AA71171	AA71171	standard; Protein; 669 AA.
ID	AA71171	standard; Protein; 669 AA.
XX	AA71171	
AC	AA71171	
XX	AA71171	
DT	21-SEP-2000	(first entry)
DE	Human Heregulin HRG-alpha protein from ORF3.	
XX	Human Heregulin HRG-alpha protein from ORF3.	
KW	Heregulin; HRG alpha; human; hair cell; inner-ear-supporting cell;	
KW	activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;	
KW	hair cell related hearing disorder; ototoxic injury; tissue damage;	
KW	acoustic assault; degenerative hearing loss; balance impairment;	
KW	treatment; surgical injury; physical injury; inner ear disorder.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PH	Key	Location/Qualifiers
FT	Cleavage-site	2..3
FT	Cleavage-site	/note= "Protease processing site"
FT	Cleavage-site	8..9
FT	Cleavage-site	/note= "Protease processing site"
FT	Cleavage-site	23..24
FT	Cleavage-site	/note= "Protease processing site"
FT	Cleavage-site	33..34
FT	Cleavage-site	/note= "Protease processing site"
FT	Cleavage-site	36..37
FT	Cleavage-site	/note= "Protease processing site"
FT	Modified-site	42..43
FT	Misc-difference	/note= "Glycosaminoglycan addition site"
FT		/label= Initiator_methionine

FT Cleavage-site 45..46 /note= "Protease processing site; cleavage results in amino terminus of HRG-alpha"

FT Cleavage-site 48..49 /note= "Protease processing site"

FT Cleavage-site 62..63 /note= "Protease processing site"

FT Modified-site 64..65 /note= "Glycosaminoglycan addition site"

FT Cleavage-site 66..67 /note= "Protease processing site"

FT Cleavage-site 86..87 /note= "Protease processing site"

FT Cleavage-site 110..111 /note= "Protease processing site"

FT Cleavage-site 123..124 /note= "Protease processing site"

FT Cleavage-site 134..135 /note= "Protease processing site"

FT Cleavage-site 142..143 /note= "Protease processing site"

FT Modified-site 151..152 /note= "Protease processing site"

FT Modified-site 164..166 /note= "Glycosaminoglycan addition site"

FT Modified-site 170..172 /note= "N-linked glycosylation site"

FT Modified-site 208..210 /note= "N-linked glycosylation site"

FT Modified-site 209..218 /note= "N-linked glycosylation site"

FT Domain 236..265 /note= "O-glycosylation site"

FT Disulfide-bond 226..240 /label= "Growth_factor_domain"

FT Disulfide-bond 234..254 /label= "HRG-alpha-GFD"

FT Disulfide-bond 256..265 /label= "Loop_1"

FT Cleavage-site 272..273 /label= "Loop_2"

FT Cleavage-site 278..279 /label= "Loop_3"

FT Cleavage-site 285..286 /note= "Protease processing site; cleavage results in carboxy terminus of HRG-alpha"

FT Cleavage-site 287..309 /note= "Protease processing site; cleavage results in carboxy terminus of HRG-alpha"

FT Domain 287..309 /label= "Transmembrane_domain"

FT Modified-site 609..611 /note= "N-linked glycosylation site"

XX WO200027426-A1.

XX 18-MAY-2000.

XX 28-OCT-1999; 99WO-US25744.

XX 07-NOV-1998; 98US-0107522.

XX (GETH) GENENTECH INC.

XX Gao W;

XX PI

XX WPI: 2000-376313/32.

XX N-PSDB; AAD00786.

XX Method for inducing hair cell generation and inner-ear-supporting cell growth regeneration and proliferation, useful for treating hearing disorders

PS Disclosure: Fig 1; 14lpp; English.

XX The patent discloses a method for inducing hair cell generation, or inner-ear-supporting cell growth, regeneration, and/or proliferation, by heregulins (HRG), the ligand for HER2/HER3 dimeric receptors. Heregulin proteins function as activators of HER-2 oncogene and result from alternate splicing of a single gene mapped to chromosome 8p. The two major types, alpha and beta HRG's are based on two variant EGF-like (epidermal growth factor) domains, that differ in their C-terminal ends. This method can be used to increase the number of inner-ear-supporting cells and for treatment of hair cell related hearing disorders and disease states associated with tissue damage, e.g. ototoxic injury, acoustic assault, degenerative hearing loss, balance impairments, damage associated with surgery or physical injury and inner ear disorders related to hair cell dysfunction. The present sequence is the human heregulin, HRG-alpha protein, derived from open reading frame (ORF) 3 of the cDNA. It is isolated from a genomic library and is used to identify other analogous HRG sequences.

CC Note: This sequence found in Figure 1 of the specification, does not match the sequence given in the sequence listing (AAY71204).

XX

SQ Sequence 669 AA;

Query Match 100.0%; Score 3470; DB 21; Length 669;
Best Local Similarity 100.0%; Pred. No. 3.7e-215;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPPIRALGPNSPAPRAVRVSVSGEMSERKEGRGKGGKK 60
|||||

DB 1 ARAPQGRSLSPSRDKLPNPPIRALGPNSPAPRAVRVSVSGEMSERKEGRGKGGKK 60
|||||

QY 61 ERSGCKKPESAGSQSPALPPRLKEMKQESAGSKVLRCETSEYSSSLRFKFKNGE 120
|||||

DB 61 ERSGCKKPESAGSQSPALPPRLKEMKQESAGSKVLRCETSEYSSSLRFKFKNGE 120
|||||

QY 121 LNRKNKPNQIKQKPKGKSELINRKASLADSGEYCKVISKLGNDASANTIVESNEII 180
|||||

DB 121 LNRKNKPNQIKQKPKGKSELINRKASLADSGEYCKVISKLGNDASANTIVESNEII 180
|||||

QY 181 TGMPASTEGAVYSSSPISRVSTEGANTSSSTSTSTTGTSHLVKCAKEKTEFCVNGEC 240
|||||

DB 181 TGMPASTEGAVYSSSPISRVSTEGANTSSSTSTSTTGTSHLVKCAKEKTEFCVNGEC 240
|||||

QY 241 FAVKDLSNPSRYLCKCQPGFTGARTENVPKMQKAEELYQKRVLTITGICIALLV 300
|||||

DB 241 FAVKDLSNPSRYLCKCQPGFTGARTENVPKMQKAEELYQKRVLTITGICIALLV 300
|||||

QY 301 GIMCVVAYCKTKQKPKGKSELINRKASLADSGEYCKVISKLGNDASANTIVESNEII 360
|||||

DB 301 GIMCVVAYCKTKQKPKGKSELINRKASLADSGEYCKVISKLGNDASANTIVESNEII 360
|||||

QY 361 VISSSEHIVERAETSFTSHYTSTAHSTTTPSHSWNGHTESILSESHVIMSSV 420
|||||

DB 361 VISSSEHIVERAETSFTSHYTSTAHSTTTPSHSWNGHTESILSESHVIMSSV 420
|||||

QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETPDSYRDSPHSERVSVAMTTPARMSP 480
|||||

DB 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETPDSYRDSPHSERVSVAMTTPARMSP 480
|||||

QY 481 VDFHTPSSPKPPSPSPVSSMTVSMVSPMEERPLLLTTPRLREKEDHHPQ 540
|||||

DB 481 VDFHTPSSPKPPSPSPVSSMTVSMVSPMEERPLLLTTPRLREKEDHHPQ 540
|||||

QY 541 QFSFHHNPADHNSLSPALRIVEDEEYETQYEPAPQVKKLANSRRRAKRTKPNCHI 600
|||||

DB 541 QFSFHHNPADHNSLSPALRIVEDEEYETQYEPAPQVKKLANSRRRAKRTKPNCHI 600
|||||

QY 601 ANRLEVDNTSSQSSNSESEDEVRGDTTFLGIONPLAASLEATPAFRADSTNPAG 660
|||||

DB 601 ANRLEVDNTSSQSSNSESEDEVRGDTTFLGIONPLAASLEATPAFRADSTNPAG 660
|||||

QY 661 RFTSQEETQ 669
|||||

Db 661 RFTQEEIQ 669

RESULT 2
AAU09886
ID AAU09886 standard; Protein; 669 AA.
AC AAU09886;
XX
XX 26-FEB-2002 (first entry)
XX Human heregulin, HRG-alpha.
XX
XX Human; heregulin; antiasthmatic; antinflammatory; vulnery; antitumor;
KW epithelial cell growth; HER2; HER3; HER4; heregulin; lung cell; asthma;
KW lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
KW chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
KW neonatal pulmonary disease; neonatal respiratory distress syndrome;
KW meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
KW acute lung injury; cystic fibrosis; surgical wound; resection;
KW growth factor; smoke inhalation; HRG-alpha.
OS Homo sapiens.
XX US2001023241-A1.
XX
XX 20-SEP-2001.
XX
XX 02-FEB-2001; 2001US-0773517.
XX
XX 04-FEB-1998; 98US-073866P.
XX 02-FEB-1999; 99US-0243198.
XX
XX (SLIW/) SLIWOWSKI M X.
XX (KERN/) KERN J A.
XX
XX Sliwowski MX, Kern JA;
XX WPI; 2001-595807/67.
XX N-PSDB; AAS18521.
XX
XX Inducing epithelial cell growth and/or proliferation, useful in the
XX treatment of respiratory disease, comprises use of heregulin ligand as
XX growth factor -
XX
XX Claim 4; Fig 1; 68pp; English.
XX
XX The invention relates to inducing epithelial cell growth and/or
XX proliferation comprising contacting a normal epithelial cell, which
XX expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
XX HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
XX heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
XX method is used for inducing epithelial cell (preferably lung cell) growth
XX and/or proliferation, for increasing lung surfactant protein A and for
XX treating respiratory distress or emphysema, for treating chronic
XX obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
XX neonatal pulmonary diseases including neonatal respiratory distress
XX syndrome, meconium aspiration syndrome, chronic lung disease of the
XX neonate, congenital diaphragmatic hernia and acute lung injuries
XX including smoke or chemical inhalation, pneumonitis due to aspiration,
XX radiation, near drowning, cystic fibrosis and other epithelial cell
XX trauma diseases including injuries associated with surgical wounds and
XX resections, ulcers, lesions and tissue tears as normal epithelial cell
XX growth factors. For treating infants/neonates with respiratory distress
XX as well as youth and adult with poor lung function due to lung injury
XX or damage. The HRG binds with varying and very high affinity to the
XX HER2, HER3 and/or HER4 receptors. The method stimulates growth and
XX proliferation of the epithelial cells, repairing and re-establishing
XX the cellular barriers of organs and allowing the affected tissue to
XX develop normal physiological functions more quickly. Hence the method
XX improves oxygenation and speeds of the development of a barrier to
XX infection while treating lung cells that are damaged by inhalation of
XX smoke resulting in emphysema. It also facilitates regeneration of

CC epithelial cells. The present sequence represents the amino acid sequence
CC of human heregulin HRG-alpha.
XX
SQ Sequence 669 AA;
Query Match 100.0%; Score 3470; DB 22; Length 669;
Best Local Similarity 100.0%; Pred. No. 3.7e-215;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ARAPQGRSLSPSRDKLPNPPIRALGNPSPAPRAVRVRSVSGMSERKEGRGKGGKK 60
Db 1 ARAPQGRSLSPSRDKLPNPPIRALGNPSPAPRAVRVRSVSGMSERKEGRGKGGKK 60
Qy 61 ERSGKKPESAGSQSPALPRLKEMKQESAAGSKLVLCETSSYSSLRFWFKNGNE 120
Db 61 ERSGKKPESAGSQSPALPRLKEMKQESAAGSKLVLCETSSYSSLRFWFKNGNE 120
Qy 121 LNRNKKQNIQKPKKSELINIKASLADSGEYCKVSKLGNDSANITIVESNEII 180
Db 121 LNRNKKQNIQKPKKSELINIKASLADSGEYCKVSKLGNDSANITIVESNEII 180
Qy 181 TGMPESTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPESTEGAYVSSSEPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
Qy 241 FMVKDLNPSRYLCKQPGFTGARTENPMKVQNEKAELYOKRVLTITIGICALLV 300
Db 241 FMVKDLNPSRYLCKQPGFTGARTENPMKVQNEKAELYOKRVLTITIGICALLV 300
Qy 301 GIMCVAYCKTKQKKLHDLRSLRERNMMNIANGPHHPNPPENVLVNOYYSKN 360
Db 301 GIMCVAYCKTKQKKLHDLRSLRERNMMNIANGPHHPNPPENVLVNOYYSKN 360
Qy 361 VISSEHIVERAETSFTSHYTSTAHSSTTQTSPHSWSNGHTESILSHSVIVMSSV 420
Db 361 VISSEHIVERAETSFTSHYTSTAHSSTTQTSPHSWSNGHTESILSHSVIVMSSV 420
Qy 421 ENSRHSSPTGGPRLNGTGGPRECNFLRHARETSDYRSPHSERYVSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRLNGTGGPRECNFLRHARETSDYRSPHSERYVSAMTTPARMSP 480
Qy 481 VDEHTSPSPKPPSEMPPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
Db 481 VDEHTSPSPKPPSEMPPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
Qy 541 QFSFHNPAHDSNLSPLRIVEDEEYETQTEPAQEPVKKLANSRRAKRTKPNNGHI 600
Db 541 QFSFHNPAHDSNLSPLRIVEDEEYETQTEPAQEPVKKLANSRRAKRTKPNNGHI 600
Qy 601 ANRLEVDNNTSSQSSNSESETEDEVEDPFLGIONPLAASLEATPAFLADSRTPAG 660
Db 601 ANRLEVDNNTSSQSSNSESETEDEVEDPFLGIONPLAASLEATPAFLADSRTPAG 660
Qy 661 RFTQEEIQ 669
Db 661 RFTQEEIQ 669
RESULT 3
AAR29570
ID AAR29570 standard; Protein; 669 AA.
XX
AC AAR29570;
XX
XX 20-APR-1993 (first entry)
XX
XX Human heregulin-alpha deduced from clone lambda gt10 her16.
XX
XX p185-HER2 ligand; her2 proto-oncogene; breast cancer; EGF receptor;
KW epidermal growth factor; tyrosine kinase-like glycoprotein;
KW HRG-alpha.
XX
XX Homo sapiens.
OS

XX Key Location/Qualifiers
 FT Domain 288..309
 FT /label= "transmembrane"
 FT /note= "probable"
 FT Region 236..265
 FT /label= EGF_motif
 FT /note= "contains 6 cysteines"
 FT Modified-site 42..43
 FT /note= "serine-glycine dipeptide potential
 glycosaminoglycan addition site -
 not part of mature HRG-alpha sequence"
 FT Modified-site 64..65
 FT /note= "serine-glycine dipeptide potential
 glycosaminoglycan addition site"
 FT Modified-site 151..152
 FT /note= "serine-glycine dipeptide potential
 glycosaminoglycan addition site"
 FT Modified-site 164..166
 FT /note= "N-linked glycosylation site"
 FT Modified-site 170..172
 FT /note= "N-linked glycosylation site"
 FT Modified-site 208..210
 FT /note= "N-linked glycosylation site"
 FT Modified-site 437..439
 FT /note= "N-linked glycosylation site"
 FT Modified-site 609..611
 FT /note= "N-linked glycosylation site"
 FT Modified-site 209..221
 FT /note= "serine-threonine potential
 O-glycosylation sites"
 FT Region 1..44
 FT /note= "deduced from ORF but initiating MET is
 at position 45 and the processed
 N-terminal residue is Ser46"
 WO9220798-A.
 XX
 PD 26-NOV-1992.
 XX
 XX 21-MAY-1992; 92WO-US04295.
 XX
 PR 24-MAY-1991; 91US-0705256.
 PR 25-SEP-1991; 91US-0765212.
 PR 08-NOV-1991; 91US-0790801.
 PR 06-MAR-1992; 92US-0847743.
 PR 11-MAY-1992; 92US-0880917.
 XX
 (GENE-) GENETECH INC.
 PI Holmes WE, Vandlen RL;
 XX
 XX WPI: 1992-415776/50.
 DR P-PSDB; AAR29570.
 XX
 PT New polypeptide(s) heregulin(s) bind the p185HER2 receptor - for
 PT diagnosing and treating tumours and screening for agonists or
 PT antagonists for binding to p185HER2
 XX
 PS Example 3; Fig 4; 139pp; English.
 XX
 CC Heregulin-alpha was isolated from the supernatant of human breast
 CC carcinoma MDA-MB-231. After purification, a major peak of tyrosine
 CC phosphorylation activity was correlated with a 45kd protein. The
 CC protein was cleaved with Lysine-C and one of the resulting peptide
 CC fragments (C-15) was used to design an oligonucleotide probe (see
 CC AAQ31547). The probe was labelled and used to screen an oligo
 CC dr-primed cDNA library constructed from human MDA-MB-231 cell mRNA
 CC in lambda gt10. Two positive clones, designated lambda gt10 her16
 CC and lambda gt10 her13, were identified and found to be identical.
 CC The sequence of lambda gt10 her16 contains a single ORF of 669
 CC amino acids. No stop codon was found in the translated sequence.
 CC See also AAQ31542-Q31548.

XX SQ Sequence 669 AA;
 Query Match 99.9%; Score 3466; DB 13; Length 669;
 Best Local Similarity 99.9%; Pred. No. 6.6e-215;
 Matches 668; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARAPORGSLSPSRDKLPNPPIRALGPNSPAPRAVVERSVSGEMSERKEGKGGK 60
 DB 1 ARAPORGSLSPSRDKLPNPPIRALGPNSPAPRAVVERSVSGEMSERKEGKGGK 60
 QY 61 ERSGKKPESAAAGSQSPALPPRLKEMKQESAAAGSKLVLCRTSEYSSSRKFKNGNE 120
 DB 61 ERSGKKPESAAAGSQSPALPPRLKEMKQESAAAGSKLVLCRTSEYSSSRKFKNGNE 120
 QY 121 LNRKNKPNQIKQKPKGKSELINAKSLADSEYKCKYISKLGNDASANITIVSNEII 180
 DB 121 LNRKNKPNQIKQKPKGKSELINAKSLADSEYKCKYISKLGNDASANITIVSNEII 180
 QY 181 TGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGBC 240
 DB 181 TGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGBC 240
 QY 241 FMVKDLSNPSRYLCKQCPGFTGARTENVMKVQNOERAEELYQKRVLTITGICIALVY 300
 DB 241 FMVKDLSNPSRYLCKQCPGFTGARTENVMKVQNOERAEELYQKRVLTITGICIALVY 300
 QY 301 GIMCVVAYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPENVQLVNOYVSKN 360
 DB 301 GIMCVVAYCKTKQKRLHDLRQSLRSENNMMNIANGPHHPNPPENVQLVNOYVSKN 360
 QY 361 VISSEHIVERAEATSFSTSHYTSTAHSTTQTSPSHSNGHTSILSESHSVIMSSV 420
 DB 361 VISSEHIVERAEATSFSTSHYTSTAHSTTQTSPSHSNGHTSILSESHSVIMSSV 420
 QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERVYAMTTPARMS 480
 DB 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYRDSPHSERVYAMTTPARMS 480
 QY 481 VDFHTPSSPKPPSEMPVSSMTYMPMAVSPFMEERPLLVTPLRLREKFDHHPQ 540
 DB 481 VDFHTPSSPKPPSEMPVSSMTYMPMAVSPFMEERPLLVTPLRLREKFDHHPQ 540
 QY 541 QFSFPHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANRRARTRPNNGHI 600
 DB 541 QFSFPHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANRRARTRPNNGHI 600
 QY 601 ANRLEVDNTSSQSSNSETEDEYVGEDTFLGTONPLAASLEATPAFRADSTNPAG 660
 DB 601 ANRLEVDNTSSQSSNSETEDEYVGEDTFLGTONPLAASLEATPAFRADSTNPAG 660
 QY 661 RFSTQEEIQ 669
 DB 661 RFSTQEEIQ 669

RESULT 4

AAR29571
 ID AAR29571 standard; Protein; 675 AA.

XX AC AAR29571;

XX DT 20-APR-1993 (first entry)

XX DE Human heregulin-beta-1 deduced from clone lambda her11.1db1.

XX KW p185-HER2 ligand; her2 proto-oncogene; breast cancer; EGF receptor;
 KW epidermal growth factor; tyrosine kinase-like glycoprotein;
 XX HRG-beta1.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

OS Homo sapiens.
PN US2001023241-A1.
XX 20-SEP-2001.
PD 02-FEB-2001; 2001US-07733517.
PF 04-FEB-1998; 98US-073866P.
PR 02-FEB-1999; 99US-0243198.
XX (SLIW/) SLINKOWSKI M X.
PA (KERN/) KERN J A.
XX Sliwowski MX, Kern JA;
PI WPI: 2001-595807/67.
XX N-PSDB; AAS18522.
DR Inducing epithelial cell growth and/or proliferation, useful in the
XX treatment of respiratory disease, comprises use of heregulin ligand as
growth factor -
Claim 4; Fig 2; 68pp; English.
CC The invention relates to inducing epithelial cell growth and/or
XX proliferation comprising contacting a normal epithelial cell, which
XX expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
XX HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
XX heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
XX method is used for inducing epithelial cell (preferably lung cell) growth
XX and/or proliferation, for increasing lung surfactant protein A and for
XX treating respiratory distress or emphysema, for treating chronic
XX obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
XX neonatal pulmonary diseases including neonatal respiratory distress
XX syndrome, meconium aspiration syndrome, chronic lung disease of the
XX neonate, congenital diaphragmatic hernia and acute lung injuries
XX including smoke or chemical inhalation, pneumonitis due to aspiration,
XX radiation, near drowning, cystic fibrosis and other epithelial cell
XX trauma diseases including injuries associated with surgical wounds and
XX resections, ulcers, lesions and tissue tears as normal epithelial cell
XX growth factors. For treating infants/neonates with respiratory distress
XX as well as youth and adult with poor lung function due to lung injury
XX or damage. The HRG binds with varying and very high affinity to the
XX HER2, HER3 and/or HER4 receptors. The method stimulates growth and
XX proliferation of the epithelial cells, repairing and re-establishing
XX the cellular barriers of organs and allowing the affected tissue to
XX develop normal physiological functions more quickly. Hence the method
XX improves oxygenation and speeds of the development of a barrier to
XX infection while treating lung cells that are damaged by inhalation of
XX smoke resulting in emphysema. It also facilitates regeneration of
XX epithelial cells. The present sequence represents the amino acid sequence
XX of human heregulin HRG-beta1.
SQ Sequence 675 AA;
Query Match 94.6%; Score 3283.5; DB 22; Length 675;
Best Local Similarity 96.7%; Pred. No. 3.8e-203;
Matches 638; Conservative 2; Mismatches 15; Indels 5; Gaps 1;
QY 15 DKLPNPPIRALGNPSPAPRAVVERSVSGEMERKGRGKGGKRGSGKKPESAAGS 74
DB 1 DKLPNPPIRALGNPSPAPRAVVERSVSGEMERKGRGKGGKRGSGKKPESAAGS 60
QY 75 QSPALPRLKEMKQBSAAGSKVLRCETSSYSSSLRPFKWFNGNELNKNKPNQIKIQK 134
DB 61 QSPALPQLKEMKQBSAAGSKVLRCETSSYSSSLRPFKWFNGNELNKNKPNQIKIQK 120
QY 135 KPGKSELRLINKASLADSGYMKVTSKLGNDGSANITIVESNEIITGMPTAGEAYVSS 194
DB 121 KPGKSELRLINKASLADSGYMKVTSKLGNDGSANITIVESNEIITGMPTAGEAYVSS 180
QY 195 ESPIRISVTEGANTSSSTTGTSHLVKCAEKTEFCVNGGECFMVKDLSNPSRYLC 254

DB 181 ESPIRISVTEGANTSSSTTGTSHLVKCAEKTEFCVNGGECFMVKDLSNPSRYLC 240
QY 255 KCPGFTGARGTENVPKVNQ-----EKAEEYLQKRVLTITGICITALLVVGIMCVVAYC 309
DB 241 KCPNEFTGDRCONVYMASFYKHLGIEFMEAEELYQKRVLTITGICITALLVVGIMCVVAYC 300
QY 310 KTKKQKRLHDLRLRQSLRSERNNMNIANGPHHPNPPEPNVQVLYVYVSKNVISSEHIVE 369
DB 301 KTKKQKRLHDLRLRQSLRSERNNMNIANGPHHPNPPEPNVQVLYVYVSKNVISSEHIVE 360
QY 370 REAETSFTSHYTSTAHTSTTVTQTPSHWSNGHTESSILSESHSVIVMSVENSRSHTSPT 429
DB 361 REAETSFTSHYTSTAHTSTTVTQTPSHWSNGHTESSILSESHSVIVMSVENSRSHTSPT 420
QY 430 GPGRLNGTGGPRECNFLRHARETDPDSYRSPHSEYVSAMTTPARMSPVDFHTPSSP 489
DB 421 GPGRLNGTGGPRECNFLRHARETDPDSYRSPHSEYVSAMTTPARMSPVDFHTPSSP 480
QY 490 KSPPEMSPPVSSMTVSMPSMAVSPFMEERPLLVTPLRLREKKEFHHPQPFSSFHNP 549
DB 481 KSPPEMSPPVSSMTVSMPSMAVSPFMEERPLLVTPLRLREKKEFHHPQPFSSFHNP 540
QY 550 AHDSNLSPLPRIVDEDEYETQYEPAQEPVKKLANSSRAKRTKPNGHIANRLEVDN 609
DB 541 AHDSNLSPLPRIVDEDEYETQYEPAQEPVKKLANSSRAKRTKPNGHIANRLEVDN 600
QY 610 TSSQSSNSESEDEVRGDDTPLGIONPLAASLEATPAFLADSTNPAGRFSTQEEIQ 669
DB 601 TSSQSSNSESEDEVRGDDTPLGIONPLAASLEATPAFLADSTNPAGRFSTQEEIQ 660
RESULT 11
AAV71202
ID AAY71202 standard; Protein; 675 AA.
XX AAY71202;
AC AAY71202;
DT 21-SEP-2000 (first entry)
XX Human Heregulin HRG-beta1 protein variant-25.
DE Heregulin; HRG beta1; human; mutant; mutein; inner-ear-supporting cell;
KW activator; HER2 receptor; HER3 receptor; HER-2 oncogene; chromosome 8p;
KW hair cell related hearing disorder; ototoxic injury; tissue damage;
KW acoustic assault; degenerative hearing loss; balance impairment;
KW treatment; surgical injury; physical injury; inner ear disorder.
XX Homo sapiens.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 31 /label= Initiator_methionine
FT FT
FT Misc-difference 253 /label= His, Trp
FT /note= "This variable site corresponds to position 223 of
the HRG-beta1 mature protein, after initiator Met (M31)."
PN WO200027426-A1.
XX 18-MAY-2000.
PD 28-OCT-1999; 99WO-US25744.
XX 07-NOV-1998; 98US-0107522.
PR (GETH) GENENTECH INC.
XX Gao W;
XX WPI: 2000-376313/32.
DR

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:44:11 ; Search time 27 Seconds
(without alignments)
729.035 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRSLSPSRDKLPN.....RLADRTNPAGRFSTQERIQ 669

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

optal number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3470	100.0	669	1	US-07-847-743B-8
2	3470	100.0	669	1	US-07-847-743B-13
3	3470	100.0	669	1	US-08-456-201-8
4	3470	100.0	669	1	US-08-456-201-13
5	3470	100.0	669	2	US-08-456-201-11
6	3470	100.0	669	2	US-08-330-161-11
7	3470	100.0	669	2	US-08-456-241-8
8	3470	100.0	669	2	US-08-456-241-13
9	3470	100.0	669	2	US-08-440-401-11
10	3470	100.0	669	2	US-08-419-878B-11
11	3470	100.0	669	4	US-09-173-480-11
12	3470	100.0	669	5	PCT-US92-04295A-8
13	3283.5	94.6	732	1	PCT-US92-04295A-13
14	3283.5	94.6	732	1	US-07-847-743B-9
15	3283.5	94.6	732	2	US-08-456-201-9
16	3283.5	94.6	732	2	US-08-456-241-9
17	3247	93.6	625	5	PCT-US92-04295A-9
18	3247	93.6	625	1	US-07-847-743B-26
19	3247	93.6	625	1	US-08-456-201-26
20	3247	93.6	625	1	US-08-456-241-26
21	3247	93.6	625	5	PCT-US92-04295A-26
22	3232.5	93.2	640	4	US-08-467-602-256
23	3227	93.0	660	4	US-08-467-602-267
24	3222.5	92.9	669	4	US-08-467-602-264
25	3213.5	92.6	687	4	US-08-467-602-273
26	3199	92.2	696	4	US-08-467-602-254
27	3193.5	92.0	707	4	US-08-467-602-268
28	3189	91.9	716	4	US-08-467-602-274
29	3149	90.7	646	4	US-08-467-602-270
30	3141.5	90.5	637	4	US-08-467-602-258
31	3133.5	90.3	645	3	US-08-753-007A-10
32	3133.5	90.3	645	4	US-09-398-496-10
33	3131.5	90.2	637	1	US-07-847-743B-28
34	3131.5	90.2	637	1	US-08-456-201-28
35	3131.5	90.2	637	2	US-08-456-241-28
36	3131.5	90.2	637	5	PCT-US92-04295A-28
37	3129.5	90.2	645	1	US-07-847-743B-27
38	3129.5	90.2	645	1	US-08-456-201-27
39	3129.5	90.2	645	1	US-08-428-926-4
40	3129.5	90.2	645	1	US-08-428-927-4
41	3129.5	90.2	645	1	US-08-428-298-4
42	3129.5	90.2	645	1	US-08-339-517-4
43	3129.5	90.2	645	2	US-08-456-241-27
44	3129.5	90.2	645	4	US-09-020-880-93
45	3129.5	90.2	645	4	US-09-101-544-93

ALIGNMENTS

RESULT 1

US-07-847-743B-8

; Sequence 8, Application US/07847743B

; Patent No. 5367060

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; TITLE OF INVENTION: Structure, Production and Use of

; TITLE OF INVENTION: Heregulin

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/847.743B

; FILING DATE: 19920306

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-May-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/765212

; FILING DATE: 25-Sep-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/790801

; FILING DATE: 08-No. 5367060-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Hensley, Max D.

; REGISTRATION NUMBER: 27,043

; REFERENCE/DOCKET NUMBER: 712P3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/266-1489

; TELEFAX: 415/952-9881

; TELEFAX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; TYPE: 669 amino acids

; TOPOLOGY: linear

; US-07-847-743B-8

Query Match

100.0%; Score 3470; DB 1; Length 669;

Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVRVERSVSGEMSERKEGKGGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVRVERSVSGEMSERKEGKGGKKK 60

QY 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCETSSYSSSLRPFKFWKNGNE 120
DB 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCETSSYSSSLRPFKFWKNGNE 120

QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 180
DB 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 180

QY 181 TGMPTAGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240

QY 241 FMYKDLNPNRYLCKQCPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300
DB 241 FMYKDLNPNRYLCKQCPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300

QY 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 360
DB 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 360

QY 361 VISSEHIVERAEATSFSTSHYTSTAHSSTVTQTPSHWSNGHTESSILSESHSVIYMSV 420
DB 361 VISSEHIVERAEATSFSTSHYTSTAHSSTVTQTPSHWSNGHTESSILSESHSVIYMSV 420

QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSDYSDSPHSRYVSAMTTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSDYSDSPHSRYVSAMTTPARMSP 480

QY 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540
DB 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540

QY 541 QFSSFHNPADHNSLSPASPLRIVEDEEYETTOEYEPAQEPVKKLANSSRAKRTKPNNGHI 600
DB 541 QFSSFHNPADHNSLSPASPLRIVEDEEYETTOEYEPAQEPVKKLANSSRAKRTKPNNGHI 600

QY 601 ANRLEVDNSTSSQSSNSETEDERVGEDTFFLGIONPLAASLEATPAFLADSRTPNAG 660
DB 601 ANRLEVDNSTSSQSSNSETEDERVGEDTFFLGIONPLAASLEATPAFLADSRTPNAG 660

QY 661 RFTQREIQ 669
DB 661 RFTQREIQ 669

RESULT 2
US-07-847-743B-13
Sequence 13, Application US/07847743B
Patent No. 5367060

GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Structure, Production and Use of
TITLE OF INVENTION: Heregulin
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080

COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/847,743B
FILING DATE: 19920306
CLASSIFICATION: 424
PRIOR APPLICATION DATA: 07/705256
FILING DATE: 24-May-1991
APPLICATION NUMBER: 07/765212
FILING DATE: 25-Sep-1991
PRIOR APPLICATION DATA: 07/790801
APPLICATION NUMBER: 08-No. 5367060-1991
FILING DATE: 08-No. 5367060-1991
ATTORNEY/AGENT INFORMATION:
NAME: Hensley, Max D.
REGISTRATION NUMBER: 27,043
REFERENCE/DOCKET NUMBER: 712P3
TELEPHONE: 415/266-1489
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-847-743B-13

Query Match 100.0%; Score 3470; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVRVERSVSGEMSERKEGKGGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNPAPRAVRVERSVSGEMSERKEGKGGKKK 60

QY 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCETSSYSSSLRPFKFWKNGNE 120
DB 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCETSSYSSSLRPFKFWKNGNE 120

QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 180
DB 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 180

QY 181 TGMPTAGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240

QY 241 FMYKDLNPNRYLCKQCPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300
DB 241 FMYKDLNPNRYLCKQCPGTGARTENPMKVQNOEKABELYQKRVLTITGICIALVY 300

QY 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 360
DB 301 GIMCVVAYCKTKQKPKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEII 360

QY 361 VISSEHIVERAEATSFSTSHYTSTAHSSTVTQTPSHWSNGHTESSILSESHSVIYMSV 420
DB 361 VISSEHIVERAEATSFSTSHYTSTAHSSTVTQTPSHWSNGHTESSILSESHSVIYMSV 420

QY 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSDYSDSPHSRYVSAMTTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETDPDSDYSDSPHSRYVSAMTTPARMSP 480

QY 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540
DB 481 VDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPPRLREKKFDHHPQ 540

QY 541 QFSSFHNPADHNSLSPASPLRIVEDEEYETTOEYEPAQEPVKKLANSSRAKRTKPNNGHI 600
DB 541 QFSSFHNPADHNSLSPASPLRIVEDEEYETTOEYEPAQEPVKKLANSSRAKRTKPNNGHI 600

QY 601 ANRLEVDNSTSSQSSNSETEDERVGEDTFFLGIONPLAASLEATPAFLADSRTPNAG 660
DB 601 ANRLEVDNSTSSQSSNSETEDERVGEDTFFLGIONPLAASLEATPAFLADSRTPNAG 660

Db 601 ANRLEVDNTSSQSSNSETEDEVRGDTFPLGIONPLAASLEATPAFLADSRTPAG 660
QY 661 RFTSQEEIQ 669
Db 661 RFTSQEEIQ 669

RESULT 3
US-08-456-201-8
; Sequence 8, Application US/08456201
; Patent No. 5641869
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145
; FILING DATE:
; APPLICATION NUMBER: 07/880,917
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; APPLICATION NUMBER: 07/765212
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/790801
; ATTORNEY/AGENT INFORMATION:
; NAME: Hensley, Max D.
; REGISTRATION NUMBER: 27,034
; REFERENCE/DOCKET NUMBER: 712P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1489
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-456-201-8

Query Match 100.0%; Score 3470; DB 1: Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPOGRSLSPSKDLFPNPIRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
Db 1 ARAPOGRSLSPSKDLFPNPIRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
QY 61 ERGSKKPESAGSOSPAPRLKEMKQESAAGSKVLRCETSEYSSSLRFKFKNGNE 120
Db 61 ERGSKKPESAGSOSPAPRLKEMKQESAAGSKVLRCETSEYSSSLRFKFKNGNE 120
QY 121 LNRKNKPNIKQKPKGSELRINKASLADSGEYMKVSKLGNDSASANTIVNESNII 180

Db 121 LNRKNKPNIKQKPKGSELRINKASLADSGEYMKVSKLGNDSASANTIVNESNII 180
QY 181 TCMPEASTGAYVSSSPIRISVSTEGANTSSTSTGTSHLVKCAEKEKTFVNGGEC 240
Db 181 TCMPEASTGAYVSSSPIRISVSTEGANTSSTSTGTSHLVKCAEKEKTFVNGGEC 240
QY 241 FMVLDLSPRYLCKCQPGFTGARTENVPKVNQOEKAEELYQKRVLTITGICIALIV 300
Db 241 FMVLDLSPRYLCKCQPGFTGARTENVPKVNQOEKAEELYQKRVLTITGICIALIV 300
QY 301 GIMCVVAYCKTKKORKKLHDLRLSLRERNMNIANGPHHPNPPENVOLVNOYVSKN 360
Db 301 GIMCVVAYCKTKKORKKLHDLRLSLRERNMNIANGPHHPNPPENVOLVNOYVSKN 360
QY 361 VISSEHIVEREAETSFSTSHYTSTAHHSTTTVTPSHSWSNGHTESILSESHSVIVMSSV 420
Db 361 VISSEHIVEREAETSFSTSHYTSTAHHSTTTVTPSHSWSNGHTESILSESHSVIVMSSV 420
QY 421 ENSRHSSPTGGPRGLNCTGGPRECNSFLRHARETPODSYRSPHSERYVSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRGLNCTGGPRECNSFLRHARETPODSYRSPHSERYVSAMTTPARMSP 480
QY 481 VDFHTPPSPKPPSPSPSPVSMVSPSMVSPSMVSPSMVSPSMVSPSMVSPSMVSPSMVSP 540
Db 481 VDFHTPPSPKPPSPSPSPVSMVSPSMVSPSMVSPSMVSPSMVSPSMVSPSMVSPSMVSP 540
QY 541 QFSSFHNPADHNSLSPASPLRIVEDEYEYETQYEPAQEPVKKLANSRRAKRTKPNNGHI 600
Db 541 QFSSFHNPADHNSLSPASPLRIVEDEYEYETQYEPAQEPVKKLANSRRAKRTKPNNGHI 600
QY 601 ANRLEVDNTSSQSSNSETEDEVRGDTFPLGIONPLAASLEATPAFLADSRTPAG 660
Db 601 ANRLEVDNTSSQSSNSETEDEVRGDTFPLGIONPLAASLEATPAFLADSRTPAG 660
QY 661 RFTSQEEIQ 669
Db 661 RFTSQEEIQ 669

RESULT 4
US-08-456-201-13
; Sequence 13, Application US/08456201
; Patent No. 5641869
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.,
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: Structure, Production and Use of
; TITLE OF INVENTION: Heregulin 2 Ligands
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,201
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/126,145
; FILING DATE:
; APPLICATION NUMBER: 07/880,917
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/765212
; APPLICATION NUMBER: 07/790801
; APPLICATION NUMBER: 07/765212

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/790801
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hensley, Max D.
;; REGISTRATION NUMBER: 27,034
;; REFERENCE/DOCKET NUMBER: 712P4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/266-1489
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-456-201-13

Query Match 100.0%; Score 3470; DB 1; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ARAPQGRSLSPSRDKLPNIPRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
1 ARAPQGRSLSPSRDKLPNIPRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
QY 61 ERSGKKPESAGSQSPALPPRLKEMKQBSAAGSKLVLCETSSYSSLRKFWKNGNE 120
61 ERSGKKPESAGSQSPALPPRLKEMKQBSAAGSKLVLCETSSYSSLRKFWKNGNE 120
QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 180
121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 180
QY 181 TGMPTGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
181 TGMPTGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQPGTGCARTENPMKVQNOEKAELQKRVLTITGICIAVV 300
241 FMVKDLSNPSRYLCKQPGTGCARTENPMKVQNOEKAELQKRVLTITGICIAVV 300
QY 301 GIMCVAYCKTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 360
301 GIMCVAYCKTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 360
QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
Db 421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYEDSPHSERYYSAMTTPARMSP 480
421 ENSRHSSPTGGPRGLNGTGGPRECNFLRHARETSDSYEDSPHSERYYSAMTTPARMSP 480
QY 481 VDPHTPSSPKPPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLVTPPRLREKKFDHPQ 540
481 VDPHTPSSPKPPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLVTPPRLREKKFDHPQ 540
QY 541 QFSFHHNPAHDSNLSLPASPLRIVEDEYETQYEPAPQEPVKKLANSRRAKRTKPNHI 600
541 QFSFHHNPAHDSNLSLPASPLRIVEDEYETQYEPAPQEPVKKLANSRRAKRTKPNHI 600
QY 601 ANRLEVDSTSSOSSNSESETEDEYEDTTPFLGIONPLAASLEATPAFLADSRNPAG 660
601 ANRLEVDSTSSOSSNSESETEDEYEDTTPFLGIONPLAASLEATPAFLADSRNPAG 660
QY 661 RFTQEQEIQ 669
661 RFTQEQEIQ 669

RESULT 5
US-08-330-161-11
; Sequence 11, Application US/08330161

;; Patent No. 5834229
;; GENERAL INFORMATION:
;; APPLICANT: Vandlen, Richard
;; APPLICANT: Holmes, William
;; TITLE OF INVENTION: Structure, Production and Use of Heregulin 2 Ligands
;; NUMBER OF SEQUENCES: 17
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/330,161
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA: 08/035430
;; APPLICATION NUMBER: 08/035430
;; FILING DATE: 22-MAR-1993
;; APPLICATION NUMBER: 07/705256
;; FILING DATE: 24-MAY-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Adler, Carolyn R.
;; REGISTRATION NUMBER: 32,324
;; REFERENCE/DOCKET NUMBER: 712C1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-2614
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; US-08-330-161-11

Query Match 100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNIPRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
1 ARAPQGRSLSPSRDKLPNIPRALGPNSPAPRAVRSVSGEMSERKEGKGKGGK 60
QY 61 ERSGKKPESAGSQSPALPPRLKEMKQBSAAGSKLVLCETSSYSSLRKFWKNGNE 120
61 ERSGKKPESAGSQSPALPPRLKEMKQBSAAGSKLVLCETSSYSSLRKFWKNGNE 120
QY 121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 180
121 LNRKNPQNTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 180
QY 181 TGMPTGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
181 TGMPTGAYVSESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FMVKDLSNPSRYLCKQPGTGCARTENPMKVQNOEKAELQKRVLTITGICIAVV 300
241 FMVKDLSNPSRYLCKQPGTGCARTENPMKVQNOEKAELQKRVLTITGICIAVV 300
QY 301 GIMCVAYCKTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 360
301 GIMCVAYCKTKQKPKSELRLINKASLADSGEYMCVISKLGNDASANITIVESNEII 360
QY 361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420
361 VISSEHIVERAEATSFSTSHYTTSTAHTSTVTTQTPSHSWSNGHTESILSESHSVIYVSSV 420

	STREET:	460 Point San Bruno Blvd	
	CITY:	South San Francisco	
	STATE:	California	
	COUNTRY:	USA	
	ZIP:	94080	
	COMPUTER READABLE FORM:		
	MEDIUM TYPE:	5.25 inch, 360 Kb floppy disk	
	COMPUTER:	IBM PC compatible	
	OPERATING SYSTEM:	PC-DOS/MS-DOS	
	SOFTWARE:	patin (Genentech)	
	CURRENT APPLICATION DATA:		
	APPLICATION NUMBER:	US/08/456,241	
	FILING DATE:	31-MAY-1995	
	CLASSIFICATION:	435	
	PRIOR APPLICATION NUMBER:	08/126145	
	FILING DATE:	23-SEP-1993	
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	07/880917	
	FILING DATE:	11-MAY-1992	
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	07/847743	
	FILING DATE:	06-MAR-1992	
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	07/790801	
	FILING DATE:	08-NOV-1991	
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	07/765212	
	FILING DATE:	25-SEP-1991	
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	07/705256	
	FILING DATE:	24-MAY-1991	
	PRIOR APPLICATION DATA:		
	ATTORNEY/AGENT INFORMATION:		
	NAME:	Lee, Wendy M.	
	REGISTRATION NUMBER:	00,000	
	REFERENCE/DOCKET NUMBER:	712P4C1D1	
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE:	415/952-9881	
	TELEFAX:	415/952-9881	
	TELEX:	910/371-7168	
	INFORMATION FOR SEQ ID NO:	13:	
	SEQUENCE CHARACTERISTICS:		
	LENGTH:	669 amino acids	
	TYPE:	amino acid	
	TOPOLOGY:	linear	
	US-08-456-241-13		
	Query Match	100.0%; Score 3470; DB 2; Length 669;	
	Best Local Similarity	100.0%; Pred. No. 1.4e-245;	
	Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ARAPOGRSLSPSRDKLFNPIRALGPNAPRAVRVERSVSGEMSERKEGKGKGGKK 60	
DB	1	ARAPOGRSLSPSRDKLFNPIRALGPNAPRAVRVERSVSGEMSERKEGKGKGGKK 60	
QY	61	ERGSGKKPSAAGSQSPPALPRRLKEMKSQESAAAGSKVLVRCETSSSEYSLSLRFKFKNGNE 120	
DB	61	ERGSGKKPSAAGSQSPPALPRRLKEMKSQESAAAGSKVLVRCETSSSEYSLSLRFKFKNGNE 120	
QY	121	LNRKNRPONIKTKPGKSELIRINKASLASDSGEYCKVIISKLGNDASANITIVESNETI 180	
DB	121	LNRKNRPONIKTKPGKSELIRINKASLASDSGEYCKVIISKLGNDASANITIVESNETI 180	
QY	181	TGPASTEGAYVSSESPIRISYSTEGANTSSSTSTGTSHLVKCAEKETFCVNGCEC 240	
DB	181	TGPASTEGAYVSSESPIRISYSTEGANTSSSTSTGTSHLVKCAEKETFCVNGCEC 240	
QY	241	FWYKDLSNPRLCYLKCPQFTGARCTENPMKVQNQKAELYQRVLTIIGICIALWV 300	
DB	241	FWYKDLSNPRLCYLKCPQFTGARCTENPMKVQNQKAELYQRVLTIIGICIALWV 300	
QY	301	GIMCVVAYCKTKQKRRKLHDLRLQSLSERNMNIANGPHHPNPENVOLVNYVSKN 360	

Query Match	100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity	100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 ARAPQRGRSLSPSRDKLFPNPRIALGNPSAPRAVRVRSVSNGEMSERKEGRGKGKKK 60
Dd	1 ARAPQRGRSLSPSRDKLFPNPRIALGNPSAPRAVRVRSVSNGEMSERKEGRGKGKKK 60
Qy	61 ERSGSKPPESAAGSQSPALPRLPKEMKSQESAAAGSKVLVRCETTSSEYSSLRPFKFKNGE 120
Dd	61 ERSGSKPPESAAGSQSPALPRLPKEMKSQESAAAGSKVLVRCETTSSEYSSLRPFKFKNGE 120
Qy	121 LNRNKKPONTKIQQKPKGSELURINKASLADSGEYMKCVISKLGNDASANITIVESNEII 180
Dd	121 LNRNKKPONTKIQQKPKGSELURINKASLADSGEYMKCVISKLGNDASANITIVESNEII 180
Qy	181 TGPMPASTEGAYVSSSEPIRISVSVEGANTSSTSTTGTSHLVKCAEKETFCVNGSEC 240
Dd	181 TGPMPASTEGAYVSSSEPIRISVSVEGANTSSTSTTGTSHLVKCAEKETFCVNGSEC 240
Qy	241 FMVKDLSNPRYLCKCPGGFTGARCTENPMKVQNQEKAELYQKRVLITIGICIALLV 300
Dd	241 FMVKDLSNPRYLCKCPGGFTGARCTENPMKVQNQEKAELYQKRVLITIGICIALLV 300
Qy	301 GMCVVAYCCTTKQRKKLHDLRLOSLSRNRMNMNIANGPHHPNPBPENVOLVNQIYSKN 360

```

; TYPE: amino acid
; TOPOLOGY: linear
US-08-440-401-11

Query Match
Best Local Similarity 100.0%; Score 3470; DB 2; Length 669;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
QY 61 ERSGKKPESAAQSOSPALPRLKEMKSOEASAGSKVLRCETSEYSSSLRFKFKNGNE 120
DB 61 ERSGKKPESAAQSOSPALPRLKEMKSOEASAGSKVLRCETSEYSSSLRFKFKNGNE 120
QY 121 LNRNKPONIKIQKPKGKSELRLINKASLADSGEYCKVISKLGNDASANITIVESNEII 180
DB 121 LNRNKPONIKIQKPKGKSELRLINKASLADSGEYCKVISKLGNDASANITIVESNEII 180
QY 181 TGMPTAGAYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGAYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FAWKDLNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICIALLV 300
DB 241 FAWKDLNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICIALLV 300
QY 301 GIMCVVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
DB 301 GIMCVVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
QY 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHSVIVMSV 420
DB 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHSVIVMSV 420
QY 421 ENSRHSSPTGGPRGLNGTGPFCNCSFLRHARETPOSYRDSPHSERVVSAMTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGPFCNCSFLRHARETPOSYRDSPHSERVVSAMTPARMSP 480
QY 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
DB 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRAKTKPNNGHI 600
DB 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRAKTKPNNGHI 600
QY 601 ANRLEVDNTSSQSSNESETEDETVGDTDFLQIQLNPLAASLEATPAFLADSTNPAG 660
DB 601 ANRLEVDNTSSQSSNESETEDETVGDTDFLQIQLNPLAASLEATPAFLADSTNPAG 660
QY 661 RFSQEEIQ 669
DB 661 RFSQEEIQ 669

```

RESULT 9

US-08-419-878B-11

Sequence 11 Application US/08419878B

Patent No 5859206

GENERAL INFORMATION:

APPLICANT: Vandlen, Richard L.

APPLICANT: Holmes, William E.

TITLE OF INVENTION: Antibodies Specific For Heregulin 2-alpha

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/419,878B
; FILING DATE: 11-Apr-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/330161
; FILING DATE: 25-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/035430
; FILING DATE: 22-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/705256
; FILING DATE: 24-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0712C2D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-419-878B-11

```

```

Query Match 100.0%; Score 3470; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
DB 1 ARAPQGRSLSPSRDKLPNPIRALGPNSPAPRAVRVRSVSGEMSERKEGRGKGKKK 60
QY 61 ERSGKKPESAAQSOSPALPRLKEMKSOEASAGSKVLRCETSEYSSSLRFKFKNGNE 120
DB 61 ERSGKKPESAAQSOSPALPRLKEMKSOEASAGSKVLRCETSEYSSSLRFKFKNGNE 120
QY 121 LNRNKPONIKIQKPKGKSELRLINKASLADSGEYCKVISKLGNDASANITIVESNEII 180
DB 121 LNRNKPONIKIQKPKGKSELRLINKASLADSGEYCKVISKLGNDASANITIVESNEII 180
QY 181 TGMPTAGAYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
DB 181 TGMPTAGAYVSSPESPIRISVSTEGANTSSSTSTSTGTSHLVKCAEKEKTCVNGGEC 240
QY 241 FAWKDLNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICIALLV 300
DB 241 FAWKDLNPSRYLCKCPGFTGARTENVPKVNQOEKAEELYOKRVLTITGICIALLV 300
QY 301 GIMCVVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
DB 301 GIMCVVAYCKTKKQKLLHDLRLQSLRSERNMNIANGPHHPNPPENVOLVQYVSKN 360
QY 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHSVIVMSV 420
DB 361 VISSEHIVERAEETSFSTSHYTSHTAHSTTTTQPSHWSNGHTESILSESHSVIVMSV 420
QY 421 ENSRHSSPTGGPRGLNGTGPFCNCSFLRHARETPOSYRDSPHSERVVSAMTPARMSP 480
DB 421 ENSRHSSPTGGPRGLNGTGPFCNCSFLRHARETPOSYRDSPHSERVVSAMTPARMSP 480
QY 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
DB 481 VDFHTPSSPKPPSEMSPVSMVSMPSMAVSPFMEERPLLVTTPRLREKFDHPQ 540
QY 541 QFSFHHNPAHDSNLSPLASPLRIVEDEYETQYEPAQEPVKKLANRRAKTKPNNGHI 600

```

Db 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAPVKKLANSRRKTKPNCHI 600
|||||

QY 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

Db 601 ANRLEVDNTSSOSSNSESETEDEYVGEDEPFLGIONPLAASLEATPAFLADSRNPAG 660
|||||

QY 661 RFTQEEIQ 669
|||||

Db 661 RFTQEEIQ 669
|||||

RESULT 10

US-09-173-480-11

; Sequence 11, Application US/09173480

; Patent No. 6399746

; GENERAL INFORMATION:

; APPLICANT: Vandlen, Richard L.

; TITLE OF INVENTION: Structure, Production and Use of

; Heregulin 2 Ligands

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: patin (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/173,480

; FILING DATE: 14-Oct-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/440,401

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/035430

; FILING DATE: 22-MAR-1993

; APPLICATION NUMBER: 07/705256

; FILING DATE: 24-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee, Wendy M.

; REGISTRATION NUMBER: 00,000

; REFERENCE/DOCKET NUMBER: 712C3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1994

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 669 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-173-480-11

Query Match 100.0%; Score 3470; DB 4; Length 669;

Best Local Similarity 100.0%; Pred. No. 1.4e-245;

Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPOGRSLSPSRDKLFNPIRALGNPAPRAVVERSVSGEMSERKEGKGGKK 60
|||||

Db 1 ARAPOGRSLSPSRDKLFNPIRALGNPAPRAVVERSVSGEMSERKEGKGGKK 60
|||||

QY 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCRTSSSYSLRKFNFKNNE 120
|||||

Db 61 ERSGGKPPSAAGSQSPALPPRLKEMKQSAAGSKLVLCRTSSSYSLRKFNFKNNE 120
|||||

;; FILING DATE: 08-NOV-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/847743
;; FILING DATE: 06-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hensley, Max D.
;; REGISTRATION NUMBER: 27,043
;; REFERENCE/DOCKET NUMBER: 712P4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-1994
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: AMINO ACIDS
;; TOPOLOGY: linear
PCT-US92-04295A-8

Query Match 100.0%; Score 3470; DB 5; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAPQGRSLSPSRDKLPFPNPIRALGPNSPAPRAVRVRSVSGEMSEKRGKGGKK 60
Db 1 ARAPQGRSLSPSRDKLPFPNPIRALGPNSPAPRAVRVRSVSGEMSEKRGKGGKK 60

Qy 61 ERSGKKPESAGSQSPALPRLKEMKQSAAGSKVLRCETSEYSSSLRKFVKNGNE 120
Db 61 ERSGKKPESAGSQSPALPRLKEMKQSAAGSKVLRCETSEYSSSLRKFVKNGNE 120

Qy 121 LNRKNKPNKIQRKPKGKSELINRKLADSGEYCKVSKLGNDSASANTIVNESNEI 180
Db 121 LNRKNKPNKIQRKPKGKSELINRKLADSGEYCKVSKLGNDSASANTIVNESNEI 180

Qy 181 TGMPTAGVSVSESPRISVSTEGANTSSSTSTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPTAGVSVSESPRISVSTEGANTSSSTSTGTSHLVKCAEKEKTCVNGGEC 240

Qy 241 FMVKDLSNPSRYLCKCPGFTGARTENPMKVONQKAEELYQKRVLTITGICIALLV 300
Db 241 FMVKDLSNPSRYLCKCPGFTGARTENPMKVONQKAEELYQKRVLTITGICIALLV 300

Qy 301 GIMCVAYCKTKKORRKLHDLRQSLRSERNMMNIANGPHPPNPENVOLVNOYVSKN 360
Db 301 GIMCVAYCKTKKORRKLHDLRQSLRSERNMMNIANGPHPPNPENVOLVNOYVSKN 360

Qy 361 VISSEHIVERAETSFSHTYHTAHSTTTVQTPSHSWNGHTESILSESHSVIVMSV 420
Db 361 VISSEHIVERAETSFSHTYHTAHSTTTVQTPSHSWNGHTESILSESHSVIVMSV 420

Qy 421 ENSRHSSPTGGPRLNGTGPRCNSFLRHARETDPDSYRSPHSERYVSAMTTPARMSP 480
Db 421 ENSRHSSPTGGPRLNGTGPRCNSFLRHARETDPDSYRSPHSERYVSAMTTPARMSP 480

Qy 481 VDFHTPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHHPQ 540
Db 481 VDFHTPSPKSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHHPQ 540

Qy 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAQEPVKLANRRRAKRPKNCHI 600
Db 541 QFSFHNPAHDSNLSPLRIVEDEYETQYEPAQEPVKLANRRRAKRPKNCHI 600

Qy 601 ANRLEVDNNTSSQSNSESETEDEVRGDDTFFLGIONPPLAASLEATPAFLADSRTPAG 660
Db 601 ANRLEVDNNTSSQSNSESETEDEVRGDDTFFLGIONPPLAASLEATPAFLADSRTPAG 660

Qy 661 RFSTQEETQ 669
Db 661 RFSTQEETQ 669

RESULT 12

PCT-US92-04295A-13
;; Sequence 13, Application PC/TUS9204295A
;; GENERAL INFORMATION:
;; APPLICANT: Genentech, Inc.
;; TITLE OF INVENTION: Structure, Production and Use of
;; TITLE OF INVENTION: Heregulin
;; NUMBER OF SEQUENCES: 30
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Genentech, Inc.
;; STREET: 460 Point San Bruno Blvd
;; CITY: South San Francisco
;; STATE: California
;; COUNTRY: USA
;; ZIP: 94080
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: patin (Genentech)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/04295A
;; FILING DATE: 19920521
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/705256
;; FILING DATE: 24-MAY-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/765212
;; FILING DATE: 25-SEP-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/790801
;; FILING DATE: 08-NOV-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/847743
;; FILING DATE: 06-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hensley, Max D.
;; REGISTRATION NUMBER: 27,043
;; REFERENCE/DOCKET NUMBER: 712P4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415/225-1994
;; TELEFAX: 415/952-9881
;; TELEX: 910/371-7168
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 669 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
PCT-US92-04295A-13

Query Match 100.0%; Score 3470; DB 5; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.4e-245;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARAPQGRSLSPSRDKLPFPNPIRALGPNSPAPRAVRVRSVSGEMSEKRGKGGKK 60
Db 1 ARAPQGRSLSPSRDKLPFPNPIRALGPNSPAPRAVRVRSVSGEMSEKRGKGGKK 60

Qy 61 ERSGKKPESAGSQSPALPRLKEMKQSAAGSKVLRCETSEYSSSLRKFVKNGNE 120
Db 61 ERSGKKPESAGSQSPALPRLKEMKQSAAGSKVLRCETSEYSSSLRKFVKNGNE 120

Qy 121 LNRKNKPNKIQRKPKGKSELINRKLADSGEYCKVSKLGNDSASANTIVNESNEI 180
Db 121 LNRKNKPNKIQRKPKGKSELINRKLADSGEYCKVSKLGNDSASANTIVNESNEI 180

Qy 181 TGMPTAGVSVSESPRISVSTEGANTSSSTSTGTSHLVKCAEKEKTCVNGGEC 240
Db 181 TGMPTAGVSVSESPRISVSTEGANTSSSTSTGTSHLVKCAEKEKTCVNGGEC 240

Qy 241 FMVKDLSNPSRYLCKCPGFTGARTENPMKVONQKAEELYQKRVLTITGICIALLV 300
Db 241 FMVKDLSNPSRYLCKCPGFTGARTENPMKVONQKAEELYQKRVLTITGICIALLV 300


```

, MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
,
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: patin (Genetech)
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/456.201
, FILING DATE:
, CLASSIFICATION: 530
, PRIOR APPLICATION DATA:
, PRIOR APPLICATION NUMBER: US/08/126,145
, FILING DATE:
, APPLICATION NUMBER: 07/880,917
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 07/705256
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 07/765212
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 07/790801
, ATTORNEY/AGENT INFORMATION:
, NAME: Hensley, Max D.
, REGISTRATION NUMBER: 27,034
, REFERENCE/DOCKET NUMBER: 712P4
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415/266-1489
, TELEFAX: 415/952-9881
, TELEX: 910/371-7168
, INFORMATION FOR SEQ ID NO: 9:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 732 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
,
, US-08-456-201-9

```

Query Match	94.6%	Score 3283.5;	DB 1;	Length 732;
Best Local Similarity	96.7%	Pred. NO. 6.7e-232;		
Matches 638; Conservative	2;	Mismatches 15;	Indels	5;

QY	15	DKLFNPNIRALGNPSNPAPRAVVRVSYSVGMSEKRGKGKGGKKEGSGKKPESAAGS	74
Db	1	DKLFNPNIRALGNPSNPAPRAVVRVSYSVGMSEKRGKGKGGKKEGSGKKPESAAGS	60
QY	75	QSPALPPRLKEMKSOESAAGSKVLRCETSEYSSLSRFKFKGNELNRKNRPONTIKIK	134
Db	61	QSPALPPQLKEMKSOESAAGSKVLRCETSEYSSLSRFKFKGNELNRKNRPONTIKIK	120
QY	135	KPGKSELIRINKASLADSGEYCKVYSKLGNDASANITIVESNEIITGHPASTEGAYVSS	194
Db	121	KPGKSELIRINKASLADSGEYCKVYSKLGNDASANITIVESNEIITGHPASTEGAYVSS	180
QY	195	ESPIRISVSTEGANTSSSTSTSTGTGTHLVKCAEKEKTCVNGGECFMVKDLSNPSRYLC	254
Db	181	ESPIRISVSTEGANTSSSTSTSTGTGTHLVKCAEKEKTCVNGGECFMVKDLSNPSRYLC	240
QY	255	KQCGFTGARTCNVPMKVQVQ-----EKAELLYQKRVLTITGICITALLVVGIMCVVAYC	309
Db	241	KCNPFZTGDRCQVYVMAFYKHLGIEFMEAFELYQKRVLTITGICITALLVVGIMCVVAYC	300
QY	310	KTKKQKKLHDLRLOSLRSERNMMANTANGPHNPPPNVQVNOVSKNVLSSEHIVE	369
Db	301	KTKKQKKLHDLRLOSLRSERNMMANTANGPHNPPPNVQVNOVSKNVLSSEHIVE	360
QY	370	REAEFSFSTSHYTSTAHHSTTVTQTPSHSWSNGHTESILSESHSVIMSVSVENSRSSPT	429
Db	361	REAEFSFSTSHYTSTAHHSTTVTQTPSHSWSNGHTESILSESHSVIMSVSVENSRSSPT	420
QY	430	GGPGRLNGTGGPRECNLSFLRHARETPDSYRDPSPHSERYVSAMTTTPARMSPVDFHTPPSP	489
Db	421	GGPGRLNGTGGPRECNLSFLRHARETPDSYRDPSPHSERYVSAMTTTPARMSPVDFHTPPSP	480
QY	490	KSPPEMSPPVSSMTVSNPMSMAVSPFMEERPLLLVTPPLRREKKFDHHPOQFSSPHNP	549
Db	481	KSPPEMSPPVSSMTVSNPMSMAVSPFMEERPLLLVTPPLRREKKFDHHPOQFSSPHNP	540

```

550 QY 550 AHDNSNLPASPLRIVEDEEYETQBYEPAQVKKLANSRRAKRTKPNGHIANRLEVDNSN 609
541 DB 541 AHDNSNLPASPLRIVEDEEYETQBYEPAQVKKLANSRRAKRTKPNGHIANRLEVDNSN 600
610 QY 610 TSSQSSNSESETEDERVGEDPFLGIONPLAASLEATPAFRLADSRNTNPAGRFSTQEEIQ 669
601 DB 601 TSSQSSNSESETEDERVGEDPFLGIONPLAASLEATPAFRLADSRNTNPAGRFSTQEEIQ 660

RESULT 15
US-08-456-241-9
; Sequence 9, Application US/08456241
; Patent No. 5840525
; GENERAL INFORMATION:
; APPLICANT: Vandlen, Richard L.
; APPLICANT: Holmes, William E.
; TITLE OF INVENTION: STRUCTURE, PRODUCTION AND USE OF
; TITLE OF INVENTION: HERGULIN
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:

```

Query Match	94.6%	Score 3283.5	DB 2	Length 732
Best Local Similarity	96.7%	Pred. No. 6.7e-232		
Matches 638	Conservative	2	Mismatches 15	Indels 5

QY 15 DKLFNPPIRALGPNPAPRAVRVERSVMSEKGRGKGKKKRGSGKKPESAAGS 74


```
Db 1 DKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGKKGSGKKPESAAGS 60
QY 75 QSPALPPRLKEMKQESAAGSKLVLRCESTSSYSSLRPKFKWFKNGNELNKNKPQNIKIQK 134
Db 61 QSPALPPQLKEMKQESAAGSKLVLRCESTSSYSSLRFKWFKNGNELNKNKPQNIKIQK 120
QY 135 KFGKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEIITGMPASTEGAYVSS 194
Db 121 KFGKSELRLINKASLADSGEYMKVISKLGNDASANITIVESNEIITGMPASTEGAYVSS 180
QY 195 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFCFWKDLNSPRLC 254
Db 181 ESPIRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTCVNGGECFCFWKDLNSPRLC 240
QY 255 KCQPGFTGARCYPNPMKYQNQ-----EKAELYOKRVLTITGICIALLVGIMCVVAYC 309
Db 241 KCPNEFTGDRCONYTWASFYKHGIEFMEAEELYOKRVLTITGICIALLVGIMCVVAYC 300
QY 310 KTKKQKKLHDLRQSLRSERNMMNIANGPHHPNPPPPENVOLVNOYVSKNVISSEHIVE 369
Db 301 KTKKQKKLHDLRQSLRSERNMMNIANGPHHPNPPPPENVOLVNOYVSKNVISSEHIVE 360
QY 370 REAETSFSSTSHYTSTAHHSTTQTTPSHWSNGHTESILSESHSVIVMSSVENSRSPT 429
Db 361 REAETSFSSTSHYTSTAHHSTTQTTPSHWSNGHTESILSESHSVIVMSSVENSRSPT 420
QY 430 GGPRGLNGTGGPRCNSFLRHARETSDSYRDSPHSERVVSAMTTPARMSPVDFHTPSSP 489
Db 421 GGPRGLNGTGGPRCNSFLRHARETSDSYRDSPHSERVVSAMTTPARMSPVDFHTPSSP 480
QY 490 KSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHHPQOFSFHHNP 549
Db 481 KSPPEMSPVSSMTVSMPSMAVSPFMEERPLLVTTPRLREKFDHHPQOFSFHHNP 540
QY 550 AHDSNLPASPLRIVEDEEYETQYEPAQEPVKKLANRRRAKTKPNGHIANRLEVDN 609
Db 541 AHDSNLPASPLRIVEDEEYETQYEPAQEPVKKLANRRRAKTKPNGHIANRLEVDN 600
QY 610 TSSQSSNSESETERVGEDTPFLGIONPLAASLEATPAFLADSRTPNAGRSTQEEIQ 669
Db 601 TSSQSSNSESETERVGEDTPFLGIONPLAASLEATPAFLADSRTPNAGRSTQEEIQ 660
```

Search completed: July 9, 2003, 10:48:41
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:45:26 ; Search time 28 Seconds
(without alignments) 2781.601 Million cell

Title: US-10-022-609-11
Perfect score: 3470
Sequence: 1 ARAPQGRGSLSPSRDKLPN.....RLADSTNPAGRESTQEEIQ 669

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 445758 seqs, 116419773 residues

total number of hits satisfying chosen parameters: 445758

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su

```

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep:*

```

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3470	100.0	669	9	US-10-022-609-11	Sequence 11, Appl
2	3470	100.0	669	10	US-09-773-517-1	Sequence 1, Appl
3	3470	100.0	669	10	US-09-792-025-1	Sequence 1, Appl
4	3470	100.0	669	10	US-09-849-868-1	Sequence 1, Appl
5	3283.5	94.6	675	10	US-09-773-517-3	Sequence 3, Appl
6	3283.5	94.6	675	10	US-09-792-025-3	Sequence 3, Appl
7	3283.5	94.6	675	10	US-09-849-868-3	Sequence 3, Appl
8	3153	90.9	632	9	US-09-946-807-16	Sequence 16, Appl
9	3153	90.9	632	10	US-09-795-668-16	Sequence 16, Appl
10	3153	90.9	632	10	US-09-795-668-16	Sequence 16, Appl
11	3133.5	90.3	645	12	US-10-096-241-10	Sequence 10, Appl
12	3131.5	90.2	637	10	US-09-773-517-5	Sequence 5, Appl
13	3131.5	90.2	637	10	US-09-792-025-5	Sequence 5, Appl
14	3131.5	90.2	637	10	US-09-849-868-5	Sequence 5, Appl
15	3047.5	87.8	629	9	US-09-946-807-14	Sequence 14, Appl
16	3047.5	87.8	629	10	US-09-795-668-14	Sequence 14, Appl
17	3047.5	87.8	629	10	US-09-795-668-14	Sequence 14, Appl
18	3039.5	87.6	637	9	US-09-946-807-13	Sequence 13, Appl
19	3039.5	87.6	637	10	US-09-795-668-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-022-609-11
: Sequence 11, Application US/10022609
: Publication No. US20030023035A1
: GENERAL INFORMATION:
: APPLICANT: Vandlen, Richard L.
: Holmes, William E.
: TITLE OF INVENTION: Heregulin 2 Ligands
: PRODUCTION AND USE OF
: PATENT:
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/022,609
: FILING DATE: 17-Dec-2001
: CLASSIFICATION: <Unknown>
: PRIORITY INFORMATION:
: APPLICATION NUMBER: US/08/440,401
: FILING DATE: 12-MAY-1995
: APPLICATION NUMBER: 08/330161
: FILING DATE: 25-OCT-1994
: APPLICATION NUMBER: 08/035430
: FILING DATE: 22-MAR-1993
: APPLICATION NUMBER: 07/705256
: FILING DATE: 24-MAY-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 712C3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881

TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-022-609-11

```
Query Match      100.0%; Score 3470; DB 9; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGKK 60
Db 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGKK 60

QY 61 ERGSGKKPESAAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
Db 61 ERGSGKKPESAAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120

QY 121 LNRKNKPQNIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
Db 121 LNRKNKPQNIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180

QY 181 TCMFASTEGAYVSSSPRIISYSTEGANTSSSTSTSTGTHLVKCAEKEKTCVNGGEC 240
Db 181 TCMFASTEGAYVSSSPRIISYSTEGANTSSSTSTSTGTHLVKCAEKEKTCVNGGEC 240

QY 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKMQVQKAEELYQKRVLTITGICITALLVY 300
Db 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKMQVQKAEELYQKRVLTITGICITALLVY 300

QY 301 GMCVVAYCKTKQKRLKDLRLQSLRSERNMMNIANGPHHPNPPENVLVNOYVSKN 360
Db 301 GMCVVAYCKTKQKRLKDLRLQSLRSERNMMNIANGPHHPNPPENVLVNOYVSKN 360

QY 361 VISSHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTSILSESHSVIWMSSV 420
Db 361 VISSHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTSILSESHSVIWMSSV 420

QY 421 ENSRHSSPPGGRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
Db 421 ENSRHSSPPGGRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480

QY 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPRLREKKFDHHPQ 540
Db 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPRLREKKFDHHPQ 540

QY 541 QFSSFHHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNGHI 600
Db 541 QFSSFHHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNGHI 600

QY 601 ANRLEVDNSTSQSSNSESETERVEDTTPGLIGNPLAASLEATPAFLADSTNPAG 660
Db 601 ANRLEVDNSTSQSSNSESETERVEDTTPGLIGNPLAASLEATPAFLADSTNPAG 660

QY 661 RSTQEEIQ 669
Db 661 RSTQEEIQ 669
```

RESULT 2
US-09-773-517-1
; Sequence 1, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-773-517-1

```
Query Match      100.0%; Score 3470; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGKK 60
Db 1 ARAPQGRSLSPSRDKLFPNPIRALGPNPAPRAVVERSVSGEMSERKEGKGKGGKK 60

QY 61 ERGSGKKPESAAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120
Db 61 ERGSGKKPESAAGSOSPALPPRLKEMKSOESAGSKLVLCRTSSEYSSLRKFWKNGNE 120

QY 121 LNRKNKPQNIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180
Db 121 LNRKNKPQNIKIQKPKGKSELINRKASLADSGEYCKVSKLGNDSASANITIVESNEII 180

QY 181 TCMFASTEGAYVSSSPRIISYSTEGANTSSSTSTSTGTHLVKCAEKEKTCVNGGEC 240
Db 181 TCMFASTEGAYVSSSPRIISYSTEGANTSSSTSTSTGTHLVKCAEKEKTCVNGGEC 240

QY 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKMQVQKAEELYQKRVLTITGICITALLVY 300
Db 241 FVVKDLSNPSRYLCKCPGFTGARTENVPKMQVQKAEELYQKRVLTITGICITALLVY 300

QY 301 GMCVVAYCKTKQKRLKDLRLQSLRSERNMMNIANGPHHPNPPENVLVNOYVSKN 360
Db 301 GMCVVAYCKTKQKRLKDLRLQSLRSERNMMNIANGPHHPNPPENVLVNOYVSKN 360

QY 361 VISSHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTSILSESHSVIWMSSV 420
Db 361 VISSHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTSILSESHSVIWMSSV 420

QY 421 ENSRHSSPPGGRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480
Db 421 ENSRHSSPPGGRGLNGTGGPRECNSFLRHARETSDSYRDSPHSERYYSAMTTARMSP 480

QY 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPRLREKKFDHHPQ 540
Db 481 VDFHTPSSPKSPPEMSPVSMTPSMAYSPFMEERPLLLVTPRLREKKFDHHPQ 540

QY 541 QFSSFHHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNGHI 600
Db 541 QFSSFHHNPAHDSNLSPLASPLRIVEDEEYETQYEPAQEPVKKLANSRRARTKPNGHI 600

QY 601 ANRLEVDNSTSQSSNSESETERVEDTTPGLIGNPLAASLEATPAFLADSTNPAG 660
Db 601 ANRLEVDNSTSQSSNSESETERVEDTTPGLIGNPLAASLEATPAFLADSTNPAG 660

QY 661 RSTQEEIQ 669
Db 661 RSTQEEIQ 669
```

RESULT 3
US-09-792-025-1
; Sequence 1, Application US/09792025
; Patent No. US20020042087A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way
City: South San Francisco
State: California
Country: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/792,025
FILING DATE: 23-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/020,598
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.
REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: P1145
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-7467
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 669 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: amino acid
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-792-025-1

Query Match 100.0%; Score 3470; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPGRSLSPSRDKLPNPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGK 60
DB 1 ARAPGRSLSPSRDKLPNPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGK 60
QY 61 ERGSKKPESAGSOSPAPLPRLKEMKSOESAGSKLVLCRTSEYSSLRKFKN 120
DB 61 ERGSKKPESAGSOSPAPLPRLKEMKSOESAGSKLVLCRTSEYSSLRKFKN 120
QY 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYCKVSKLGNDSANITIVES 180
DB 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYCKVSKLGNDSANITIVES 180
QY 181 TCMPTAGAYVSSSPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCV 240
DB 181 TCMPTAGAYVSSSPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCV 240
QY 241 FMVKDLSNPSRYLCKCPGFTGARTENVPKVNQERAEELYQKRVLTITGIC 300
DB 241 FMVKDLSNPSRYLCKCPGFTGARTENVPKVNQERAEELYQKRVLTITGIC 300
QY 301 GIMCVYAYCKTKKQKRLHDLRQSLRSERNMMNIANGPHHPNPPNPVQV 360
DB 301 GIMCVYAYCKTKKQKRLHDLRQSLRSERNMMNIANGPHHPNPPNPVQV 360
QY 361 VISSHHIVERAEATSFSTSHYTSHTAHSTTVTQTPSHWSNGHTSILSES 420
DB 361 VISSHHIVERAEATSFSTSHYTSHTAHSTTVTQTPSHWSNGHTSILSES 420
QY 421 ENSRHSSTGGPRGLNGTGGPRECNFLRHARETSDYRDSHPHSEYVSAMT 480
DB 421 ENSRHSSTGGPRGLNGTGGPRECNFLRHARETSDYRDSHPHSEYVSAMT 480
QY 481 VDFHTPSPKPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 540
DB 481 VDFHTPSPKPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 540

QY 541 QFSFHHNPAHDSNLSPLRIVEDEEYETQYEPAQEPVKKLANRRARTRKPN 600
DB 541 QFSFHHNPAHDSNLSPLRIVEDEEYETQYEPAQEPVKKLANRRARTRKPN 600
QY 601 ANRLEVDNSTSSQSSSESETEDEVRGDTPLGIONPLAASLEATPAFLAD 660
DB 601 ANRLEVDNSTSSQSSSESETEDEVRGDTPLGIONPLAASLEATPAFLAD 660
QY 661 RFSTQEEIQ 669
DB 661 RFSTQEEIQ 669

RESULT 4
US-09-849-868-1
; Sequence 1, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-849-868-1

Query Match 100.0%; Score 3470; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 4e-173;
Matches 669; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAPGRSLSPSRDKLPNPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGK 60
DB 1 ARAPGRSLSPSRDKLPNPIRALGPNSPAPRAVVRVSVSGEMSEKRGKGGK 60
QY 61 ERGSKKPESAGSOSPAPLPRLKEMKSOESAGSKLVLCRTSEYSSLRKFKN 120
DB 61 ERGSKKPESAGSOSPAPLPRLKEMKSOESAGSKLVLCRTSEYSSLRKFKN 120
QY 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYCKVSKLGNDSANITIVES 180
DB 121 LNRKNKPNQIKQKPKGKSELINKASLADSGEYCKVSKLGNDSANITIVES 180
QY 181 TCMPTAGAYVSSSPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCV 240
DB 181 TCMPTAGAYVSSSPRISVSTEGANTSSSTSTSTTGTSHLVKCAEKEKTCV 240
QY 241 FMVKDLSNPSRYLCKCPGFTGARTENVPKVNQERAEELYQKRVLTITGIC 300
DB 241 FMVKDLSNPSRYLCKCPGFTGARTENVPKVNQERAEELYQKRVLTITGIC 300
QY 301 GIMCVYAYCKTKKQKRLHDLRQSLRSERNMMNIANGPHHPNPPNPVQV 360
DB 301 GIMCVYAYCKTKKQKRLHDLRQSLRSERNMMNIANGPHHPNPPNPVQV 360
QY 361 VISSHHIVERAEATSFSTSHYTSHTAHSTTVTQTPSHWSNGHTSILSES 420
DB 361 VISSHHIVERAEATSFSTSHYTSHTAHSTTVTQTPSHWSNGHTSILSES 420
QY 421 ENSRHSSTGGPRGLNGTGGPRECNFLRHARETSDYRDSHPHSEYVSAMT 480
DB 421 ENSRHSSTGGPRGLNGTGGPRECNFLRHARETSDYRDSHPHSEYVSAMT 480
QY 481 VDFHTPSPKPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 540
DB 481 VDFHTPSPKPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 540

Db 301 NVOLVNOYVSKNVISSEHIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360
QY 409 SESHVIVMSSVNSRHSSPTGGPRLNGTGGPRECNFLRHARETTPDSYRDSPHSERY 468
Db 361 SESHVIVMSSVNSRHSSPTGGPRLNGTGGPRECNFLRHARETTPDSYRDSPHS - -Y 418
QY 469 VSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSFPMEERPLLVLTPP 528
Db 419 VSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSFPMEERPLLVLTPP 478
QY 529 RLREKFDHHPQOFSFHHNPAHDSNLSLPAFLRIVEDEYETQYEPYPAQEPVKKLANS 588
Db 479 RLREKFDHHPQOFSFHHNPAHDSNLSL -DLRIVEDEYETQYEPYPAQEPVKKLANS 536
QY 589 RRAKRTKPNGHIANRLVDSNTSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLVDSNTSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPA 596
QY 649 FRLADSRTPNAGRFSTQEEIQ 669
Db 597 FRLADSRTPNAGRFSTQEEIQ 617

RESULT 9

US-09-795-668-16
; Sequence 16, Application US/09795668
; Patent No. US2002004557A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-668-16

Query Match 90.9%; Score 3153; DB 10; Length 632;
Best Local Similarity 98.7%; Pred. No. 1.2e-156;
Matches 613; Conservative 0; Mismatches 2; Indels 6; Gaps 3;

QY 49 KEGRGKGGKKGKGGKPPESAGSQSPALPRLKEMKSOESAAGSKVLVRCETSEYS 108
Db 3 KEGRGKGGKKGKGGKPPESAGSQSPALPRLKEMKSOESAAGSKVLVRCETSEYS 62
QY 109 SLRFKFWKNGNENLNKRNKQNIQKPKGSELINKASLADSGEYMKVCKVISKLGNDAS 168
Db 63 SLRFKFWKNGNENLNKRNKQNIQKPKGSELINKASLADSGEYMKVCKVISKLGNDAS 122
QY 169 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAE 228
Db 123 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAE 182
QY 229 KERTFCVNGEGCFMWKDLNPSRYLCKCPGFTGARTENVPKMQNQKAEELYOKRYL 288
Db 183 KERTFCVNGEGCFMWKDLNPSRYLCKCPGFTGARTENVPKMQNQKAEELYOKRYL 242
QY 289 TITGICALLVGVIMCVAYCKTKKQKKLHDLRLQSLRSERNMNIANGPHHPNPPPE 348
Db 243 TITGICALLVGVIMCVAYCKTKKQKKLHDLRLQSLRS -NNTNANGPHHPNPPPE 300
QY 349 NVOLVNOYVSKNVISSEHIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 408
Db 301 NVOLVNOYVSKNVISSEHIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360

QY 409 SESHVIVMSSVNSRHSSPTGGPRLNGTGGPRECNFLRHARETTPDSYRDSPHSERY 468
Db 361 SESHVIVMSSVNSRHSSPTGGPRLNGTGGPRECNFLRHARETTPDSYRDSPHS - -Y 418
QY 469 VSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSFPMEERPLLVLTPP 528
Db 419 VSAMTTPARMSPVDFHTPPSPKSPPEMSPVSSMTVSMPSMAVSFPMEERPLLVLTPP 478
QY 529 RLREKFDHHPQOFSFHHNPAHDSNLSLPAFLRIVEDEYETQYEPYPAQEPVKKLANS 588
Db 479 RLREKFDHHPQOFSFHHNPAHDSNLSL -DLRIVEDEYETQYEPYPAQEPVKKLANS 536
QY 589 RRAKRTKPNGHIANRLVDSNTSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLVDSNTSSQSSNSETEDEVRGDDTFLGIONPLAASLEATPA 596
QY 649 FRLADSRTPNAGRFSTQEEIQ 669
Db 597 FRLADSRTPNAGRFSTQEEIQ 617

RESULT 10

US-09-795-686-16
; Sequence 16, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09795,686
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-795-686-16

Query Match 90.9%; Score 3153; DB 10; Length 632;
Best Local Similarity 98.7%; Pred. No. 1.2e-156;
Matches 613; Conservative 0; Mismatches 2; Indels 6; Gaps 3;

QY 49 KEGRGKGGKKGKGGKPPESAGSQSPALPRLKEMKSOESAAGSKVLVRCETSEYS 108
Db 3 KEGRGKGGKKGKGGKPPESAGSQSPALPRLKEMKSOESAAGSKVLVRCETSEYS 62
QY 109 SLRFKFWKNGNENLNKRNKQNIQKPKGSELINKASLADSGEYMKVCKVISKLGNDAS 168
Db 63 SLRFKFWKNGNENLNKRNKQNIQKPKGSELINKASLADSGEYMKVCKVISKLGNDAS 122
QY 169 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAE 228
Db 123 ANTIIVESNEITGMPASTEGAYVSSSEPIRISVSTEGANTSSSTSTSTGTSHLVKCAE 182
QY 229 KERTFCVNGEGCFMWKDLNPSRYLCKCPGFTGARTENVPKMQNQKAEELYOKRYL 288
Db 183 KERTFCVNGEGCFMWKDLNPSRYLCKCPGFTGARTENVPKMQNQKAEELYOKRYL 242
QY 289 TITGICALLVGVIMCVAYCKTKKQKKLHDLRLQSLRSERNMNIANGPHHPNPPPE 348
Db 243 TITGICALLVGVIMCVAYCKTKKQKKLHDLRLQSLRS -NNTNANGPHHPNPPPE 300
QY 349 NVOLVNOYVSKNVISSEHIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 408
Db 301 NVOLVNOYVSKNVISSEHIVERAEATSFSTSHYSTAHHSTVTTQTPSHWSNGHTESIL 360
QY 409 SESHVIVMSSVNSRHSSPTGGPRLNGTGGPRECNFLRHARETTPDSYRDSPHSERY 468

Db 361 SESHSVIVSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETDPDSTRDPSHS--Y 418
Qy 469 VSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLLVTPP 528
Db 419 VSAMTTPARMSPVDFHTPSSPKSPSPSEMSPVSSMTVSMPSMAVSPFMEERPLLLVTPP 478
Qy 529 RLREKKFDHHPQOFSFHHNPAHNSLIPASPLRIVEDEEYETTOEYEPAPQEPVKKLANS 588
Db 479 RLREKKFDHHPQOFSFHHNPAHNSLIP--DLRIVEDEEYETTOEYEPAPQEPVKKLANS 536
Qy 589 RRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDERVGEDTFLGTONPLAASLEATPA 648
Db 537 RRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDERVGEDTFLGTONPLAASLEATPA 596
Qy 649 FRLADSRTPNAGRFSTQBEIQ 669
Db 597 FRLADSRTPNAGRFSTQBEIQ 617

RESULT 11

US-10-096-241-10

Sequence 10, Application US/10096241

; Patent No. US20020127594A1

; GENERAL INFORMATION:

; APPLICANT: Gearing, David P.

; TITLE OF INVENTION: DON-1 GENE AND POLYPEPTIDES

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/096,241

; FILING DATE: 12-Mar-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/699,591

; FILING DATE: 19-AUG-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Fasse, J. Peter

; REGISTRATION NUMBER: 32,983

; REFERENCE/DOCKET NUMBER: 07334/022001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 645 amino acids

; TYPE: amino acid

; STRANDEDNESS: not relevant

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-096-241-10

Query Match 90.3%; Score 3133.5; DB 12; Length 645;

Best Local Similarity 96.7%; Pred. No. 1.2e-155;

Matches 609; Conservative 1; Mismatches 15; Indels 5; Gaps 1;

Qy 45 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 104

Db 1 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 60

Qy 45 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 104

Db 1 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 60

Qy 105 SEYSSLRFKWFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 164
Db 61 SEYSSLRFKWFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120
Qy 165 DSANATIVESNEIITGMPASTEGAVYSSSESPRISVSTEGANTSSSTSTGTGSHLV 224
Db 121 DSANATIVESNEIITGMPASTEGAVYSSSESPRISVSTEGANTSSSTSTGTGSHLV 180
Qy 225 KCAKEKTEFCVNGGCEFMVLDLSPSRVLCQPGFTGARTENVPKVNQ-----EKA 279
Db 181 KCAKEKTEFCVNGGCEFMVLDLSPSRVLCQPGFTGARTENVPKVNQ-----EKA 240
Qy 280 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKKQKRLHDLRQLSRERNNMNIANG 339
Db 241 EELYQKRVLTITGICIAALLVVGIMCVVAYCKTKKQKRLHDLRQLSRERNNMNIANG 300
Qy 340 PHHPNPPENVQLVNYVSKNVISSEHIVREARETSESTSHYTSTAHHSTVTTQTPSHSW 399
Db 301 PHHPNPPENVQLVNYVSKNVISSEHIVREARETSESTSHYTSTAHHSTVTTQTPSHSW 360
Qy 400 SNHGTESILSESHSVIVMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSY 459
Db 361 SNHGTESILSESHSVIVMSSVNSRHSSPTGGPRGLNGTGGPRECNSFLRHARETSDSY 420
Qy 460 RDSPHSERVYSAMTTPARMSPVDPHTPSSPKSPSPSEMSPVSSMTVSMPSMAVSPFMEEE 519
Db 421 RDSPHSERVYSAMTTPARMSPVDPHTPSSPKSPSPSEMSPVSSMTVSMPSMAVSPFMEEE 480
Qy 520 RPLLLVTPRLREKKFDHHPQOFSFHHNPAHNSLIPASPLRIVEDEEYETTOEYEPAPQ 579
Db 481 RPLLLVTPRLREKKFDHHPQOFSFHHNPAHNSLIPASPLRIVEDEEYETTOEYEPAPQ 540
Qy 580 EPVKKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDERVGEDTFLGIONPL 639
Db 541 EPVKKLANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDERVGEDTFLGIONPL 600
Qy 640 AASLEATPAPRLADSRTPNAGRFSTQBEIQ 669
Db 601 AASLEATPAPRLADSRTPNAGRFSTQBEIQ 630

RESULT 12

US-09-773-517-5

; Sequence 5, Application US/09773517

; Patent No. US20010023241A1

; GENERAL INFORMATION:

; APPLICANT: Sliwkowski, Mark

; APPLICANT: Kern, Jeffrey

; TITLE OF INVENTION: Use of Heregulin as a Growth Factor

; FILE REFERENCE: P1145R1

; CURRENT APPLICATION NUMBER: US/09/773,517

; CURRENT FILING DATE: 2001-02-02

; PRIOR APPLICATION NUMBER: 09/243,198

; PRIOR FILING DATE: 1999-02-02

; NUMBER OF SEQ ID NOS: 14

; SEQ ID NO 5

; LENGTH: 637

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-773-517-5

Query Match 90.2%; Score 3131.5; DB 10; Length 637;

Best Local Similarity 97.3%; Pred. No. 1.5e-155;

Matches 608; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

Qy 45 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 104

Db 1 MSERKEGRGKGKKGKRGSGKPPESAGSQSPALPRLKEMKSQESAAGSKLVLRCTS 60

Qy 105 SEYSSLRFKWFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 164

Db 61 SEYSSLRFKWFKNGNLRNKNKPNQIKQKPGKSELINKASLADSGEYMKVSKLGN 120

QY 165 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTSTGTSHLV 224
121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTSTGTSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCENYPMKVQNEKAEELYQ 284
181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCENYPMKVQNEKAEELYQ 237
QY 285 KVLVITGICIAALLVVGIMCVVAYCTKKQKRLDRLQSLRSENNMNIANGPHHN 344
238 KVLVITGICIAALLVVGIMCVVAYCTKKQKRLDRLQSLRSENNMNIANGPHHN 297
QY 345 PPNENVOLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSNGHT 404
298 PPNENVOLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSNGHT 357
QY 405 ESILSESHSVIVSSVENSRRHSPGGRGLNGTGGPRECHSFLRHARETSDSYRDSH 464
358 ESILSESHSVIVSSVENSRRHSPGGRGLNGTGGPRECHSFLRHARETSDSYRDSH 417
465 SERYVSAMTTPARMSVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 524
418 SERYVSAMTTPARMSVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 477
QY 525 VTPPRLREKFKDHPHQFSSFHNNPAHDSNLSLPASPLRIVEDEYETTQYEPAQEPVK 584
478 VTPPRLREKFKDHPHQFSSFHNNPAHDSNLSLPASPLRIVEDEYETTQYEPAQEPVK 537
QY 585 LANSRAKRTKPNHIANLEVDNSTSSQSSSESETEDEYEDTTPFLGIONPLAASLE 644
538 LANSRAKRTKPNHIANLEVDNSTSSQSSSESETEDEYEDTTPFLGIONPLAASLE 597
QY 645 ATPAFELADSRTPNAGRFSTQEEIQ 669
598 ATPAFELADSRTPNAGRFSTQEEIQ 622

RESULT 13

US-09-792-025-5
Sequence 5, Application US/09792025
Patent No. US20020042087A1
GENERAL INFORMATION:
APPLICANT: Sliwowski, Mark
Kern, Jeffrey
TITLE OF INVENTION: Use of Heregulin as a Growth Factor
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Genentech, Inc.
STREET: 1 DNA Way
City: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/792,025
FILING DATE: 23-Feb-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/020,598
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Schwartz, Timothy R.
REGISTRATION NUMBER: 32171
REFERENCE/DOCKET NUMBER: P1145
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-7467
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
MOLECULE TYPE: amino acid
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-792-025-5

Query Match 90.2%; Score 3131.5; DB 10; Length 637;
Best Local Similarity 97.3%; Pred. No. 1.5e-155;
Matches 608; Conservative 3; Mismatches 11; Indels 3; Gaps 1;
QY 45 MSEREGRGKGGKKGKGGKPPSAAGSQSPALPPRLKEMKQESAAAGSKVLRCETS 104
1 MSEREGRGKGGKKGKGGKPPSAAGSQSPALPPRLKEMKQESAAAGSKVLRCETS 60
QY 105 SEYSLRFKWFNGNGLNKKPNQIKIQQKPKGSELRLINKASLADSGYMKVSKLGN 164
61 SEYSLRFKWFNGNGLNKKPNQIKIQQKPKGSELRLINKASLADSGYMKVSKLGN 120
QY 165 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTGTSHLV 224
121 DSANITIVESNEIITGMPASTEGAYVSSSPIRISVSTEGANTSSSTSTGTSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCENYPMKVQNEKAEELYQ 284
181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQPGFTGARCENYPMKVQNEKAEELYQ 237
QY 285 KVLVITGICIAALLVVGIMCVVAYCTKKQKRLDRLQSLRSENNMNIANGPHHN 344
238 KVLVITGICIAALLVVGIMCVVAYCTKKQKRLDRLQSLRSENNMNIANGPHHN 297
QY 345 PPNENVOLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSNGHT 404
298 PPNENVOLVQVSKNVISSEHIVERAEATSFSTSHYTSTAHHSTVTTQTPSHSNGHT 357
QY 405 ESILSESHSVIVSSVENSRRHSPGGRGLNGTGGPRECHSFLRHARETSDSYRDSH 464
358 ESILSESHSVIVSSVENSRRHSPGGRGLNGTGGPRECHSFLRHARETSDSYRDSH 417
QY 465 SERYVSAMTTPARMSVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 524
418 SERYVSAMTTPARMSVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 477
QY 525 VTPPRLREKFKDHPHQFSSFHNNPAHDSNLSLPASPLRIVEDEYETTQYEPAQEPVK 584
478 VTPPRLREKFKDHPHQFSSFHNNPAHDSNLSLPASPLRIVEDEYETTQYEPAQEPVK 537
QY 585 LANSRAKRTKPNHIANLEVDNSTSSQSSSESETEDEYEDTTPFLGIONPLAASLE 644
538 LANSRAKRTKPNHIANLEVDNSTSSQSSSESETEDEYEDTTPFLGIONPLAASLE 597
QY 645 ATPAFELADSRTPNAGRFSTQEEIQ 669
598 ATPAFELADSRTPNAGRFSTQEEIQ 622

RESULT 14

US-09-849-868-5
Sequence 5, Application US/09849868
Patent No. US20020081299A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Gao, Wei-Qiang
TITLE OF INVENTION: HAIR CELL DISORDERS
FILE REFERENCE: GENENT.035VPC
CURRENT APPLICATION NUMBER: US/09/849,868
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 637

TYPE: PRT
ORGANISM: Homo sapiens
US-09-849-868-5

Query Match 90.2%; Score 3131.5; DB 10; Length 637;
Best Local Similarity 97.3%; Pred. No. 1.5e-155;
Matches 608; Conservative 3; Mismatches 11; Indels 3; Gaps 1;

QY 45 MSERKGRGKGGKKGKSPALPPRLKEMKQSAAGSKLVLRCEFS 104
DB 1 MSERKGRGKGGKKGKSPALPPRLKEMKQSAAGSKLVLRCEFS 60
QY 105 SEYSSLRFWKNGNLRNKNPQNIQKPKGKSELRLNKASLADSGYMKVISKLN 164
DB 61 SEYSSLRFWKNGNLRNKNPQNIQKPKGKSELRLNKASLADSGYMKVISKLN 120
QY 165 DSANITIVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 224
DB 121 DSANITIVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLV 180
QY 225 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKQPGFTGARTENPMKVQNOKEABELYQ 284
DB 181 KCAEKEKTCVNGGECFMVKDLSNPSRYLCKQPGFTGARTENPMKVQNOKEABELYQ 237
QY 285 KEVLRTGICIALLVGIMCVAYCKTKQKRLHDLRLQSLRSERNMMNIANGPHHPN 344
DB 238 KEVLRTGICIALLVGIMCVAYCKTKQKRLHDLRLQSLRSERNMMNIANGPHHPN 297
QY 345 PPPENQVLYQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHT 404
DB 298 PPPENQVLYQVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHT 357
QY 405 ESILSESHVIVYVSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPH 464
DB 358 ESILSESHVIVYVSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPH 417
QY 465 SERYSAMTPARMSPVDFTHTSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 524
DB 418 SERYSAMTPARMSPVDFTHTSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLL 477
QY 525 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQEPVKK 584
DB 478 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQEPVKK 537
QY 585 LANSRAKTKPNGHIANRLEVDNSTSSQSSNSESETEDEYVGEDTPFLGIONPLAASLE 644
DB 538 LANSRAKTKPNGHIANRLEVDNSTSSQSSNSESETEDEYVGEDTPFLGIONPLAASLE 597
QY 645 ATPAFRLADSRTPNAGRFSTQEEIQ 669
DB 598 ATPAFRLADSRTPNAGRFSTQEEIQ 622

RESULT 15
US-09-946-807-14
Sequence 14, Application US/09946807
Patent No. US20020165144A1
GENERAL INFORMATION:
APPLICANT: Stefanesson, Hreinn
APPLICANT: Steinhorsdottir, Valgerdur
APPLICANT: Gulcher, Jeffrey R.
TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
FILE REFERENCE: 2345.2004-001
CURRENT APPLICATION NUMBER: US/09/946,807
CURRENT FILING DATE: 2001-09-05
PRIOR APPLICATION NUMBER: US/09/795,668
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 09/515,716
PRIOR FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 1531
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 629

TYPE: PRT
ORGANISM: Homo sapiens
US-09-946-807-14

Query Match 87.8%; Score 3047.5; DB 9; Length 629;
Best Local Similarity 96.3%; Pred. No. 3.6e-151;
Matches 598; Conservative 2; Mismatches 12; Indels 9; Gaps 4;

QY 49 KEGRGKGGKKGKKGKSPALPPRLKEMKQSAAGSKLVLRCEFS 108
DB 3 KEGRGKGGKKGKKGKSPALPPRLKEMKQSAAGSKLVLRCEFS 62
QY 109 SLRPFWFNNGNLRNKNPQNIQKPKGKSELRLNKASLADSGYMKVISKLNDGAS 168
DB 63 SLRPFWFNNGNLRNKNPQNIQKPKGKSELRLNKASLADSGYMKVISKLNDGAS 122
QY 169 ANITIVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLVKCAE 228
DB 123 ANITIVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTTGTSHLVKCAE 182
QY 229 KECTFCVNGGECFMVKDLSNPSRYLCKQPGFTGARTENPMKVQNOKEABELYQ 288
DB 183 KECTFCVNGGECFMVKDLSNPSRYLCKQPGFTGARTENPMKVQNOKEABELYQ 239
QY 289 TITGICIALLVGIMCVAYCKTKQKRLHDLRLQSLRSERNMMNIANGPHHPNPPPE 348
DB 240 TITGICIALLVGIMCVAYCKTKQKRLHDLRLQSLRSERNMMNIANGPHHPNPPPE 297
QY 349 NVOLVQVYVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTESIL 408
DB 298 NVOLVQVYVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSWSNGHTESIL 357
QY 409 SESHVIVYVSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRY 468
DB 358 SESHVIVYVSSVENSRRHSPTGGPRGLNGTGGPRECSFLRHARETDPDYSRDSPHSRY 415
QY 469 VSAMTPARMSPVDFTHTSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPP 528
DB 416 VSAMTPARMSPVDFTHTSSPKSPSEMSPPVSSMTVSMPSMAVSPFMEERPLLLVTPP 475
QY 529 RLREKKFDHHPQOFSFHHNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQEPVKKLANS 588
DB 476 RLREKKFDHHPQOFSFHHNPAHDSNLSLPASPLRIVEDEEYETTQYEPAQEPVKKLANS 533
QY 589 RRAKTKPNGHIANRLEVDNSTSSQSSNSESETEDEYVGEDTPFLGIONPLAASLEATPA 648
DB 534 RRAKTKPNGHIANRLEVDNSTSSQSSNSESETEDEYVGEDTPFLGIONPLAASLEATPA 593
QY 649 FRLADSRTPNAGRFSTQEEIQ 669
DB 594 FRLADSRTPNAGRFSTQEEIQ 614

Search completed: July 9, 2003, 10:49:15
Job time : 30 secs

Result No.	Score	Query %			DB	ID	Description
		Match	Length	Match			
1	3243	93.5	640	2	A43273	heregulin precursor	
2	3135.5	90.4	637	2	A43273	heregulin precursor	
3	3133.5	90.3	645	2	B43273	heregulin, splice	
4	3021.5	87.1	639	2	I61719	neu differentiation	
5	2922	84.2	636	2	I61718	neu differentiation	
6	2917	84.1	662	2	I61722	neu differentiation	
7	2314	66.7	602	2	A45769	acetylcholine rece	
8	2204	63.5	462	2	I38404	neu differentiation	
9	1804	52.0	350	2	I38403	neu differentiation	
10	1109	32.0	241	2	D43273	heregulin precursor	
11	1024	29.5	241	2	S32359	glial growth facto	
12	1013.5	29.2	850	2	JC5700	ErbB kinase activa	
13	1012.5	29.2	860	2	JC5702	ErbB kinase activa	
14	1001.5	28.9	868	2	JC5701	ErbB kinase activa	
15	985	28.4	230	2	A36210	neu differentiation	
16	750.5	21.6	422	2	S32357	glial growth facto	
17	710	20.5	175	2	I38408	neu differentiation	
18	598.5	17.2	125	2	S62676	heregulin isoform	
19	581	16.7	125	2	I38405	neu differentiation	
20	301	8.7	713	2	T44447	neuregulin-3 (lipo	
21	273	7.9	296	2	A56943	sensory/motor neur	
22	180.5	5.2	5327	2	T13564	microtubule-associ	
23	179.5	5.2	2254	2	D86215	protein T6D22.14 [
24	176	5.1	734	2	B42680	nucleolus-cytoplas	
25	172	5.0	3507	2	T34513	hypothetical prote	
26	170	4.9	990	2	I51618	nucleolar phosphop	
27	166.5	4.8	1459	2	T32771	hypothetical prote	
28	166.5	4.8	1510	2	T33100	hypothetical prote	
29	165	4.8	1091	1	I3CHNL	neuronal cell adhes	

QY 165 DSANANITIVESNEIITGMPASTEGAYVSSSEPIRISYSTEGANTSSSTSTSTTGTSHLV 224
DB 121 DSANANITIVESNEIITGMPASTEGAYVSSSEPIRISYSTEGANTSSSTSTSTTGTSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARCCTENPMKVONOEKAEELYQ 284
DB 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARCCTENPMKVONOEKAEELYQ 240
QY 285 KRVLTITGICIAALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNMIANGPHHPN 344
DB 241 KRVLTITGICIAALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNMIANGPHHPN 300
QY 345 PPNENVOLVQVSVKNVSSSEHIVERAEATSESTSHYTSHTAHSTVTTQTPSHSWSNGHT 404
DB 301 PPNENVOLVQVSVKNVSSSEHIVERAEATSESTSHYTSHTAHSTVTTQTPSHSWSNGHT 360
QY 405 ESILSESHSVIYVSSVENSRRHSSPTGGPRGLNGTGGPRECNFSLRHARETSDSYRDSPH 464
DB 361 ESILSESHSVIYVSSVENSRRHSSPTGGPRGLNGTGGPRECNFSLRHARETSDSYRDSPH 420
QY 465 SERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAYSPFMEERPLLL 524
DB 421 SERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAYSPFMEERPLLL 480
QY 525 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETQVEYPAQEPVK 584
DB 481 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETQVEYPAQEPVK 540
QY 585 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDETVGDTPLFIGNPLAASLE 644
DB 541 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDETVGDTPLFIGNPLAASLE 600
QY 645 ATPAFRLADSRTPNAGRFSTQEEIQ 669
DB 601 ATPAFRLADSRTPNAGRFSTQEEIQ 625
RESULT 2
C43273
heregulin precursor, splice form beta-2 - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C:Accession: C43273; 138407
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yansu
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: C43273
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-637 <HOL>
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: 138407
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 119-406 <RES>
A:Cross-references: EMBL:U02329; NID:g408408; PIDN:AAA19954.1; PID:g408409
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology
C:Keywords: alternative splicing
F:182-221/Domain: EGF homology <EGF>
Query Match 90.4%; Score 3135.5; DB 2; Length 637;
Best Local Similarity 97.4%; Pred. No. 2.2e-167;
Matches 609; Conservative 2; Mismatches 11; Indels 3; Gaps 1;
QY 45 MSEREGRGKGGKKERGGKPPSAAGSQSPALPPLKEMKSOESAGSKLVLRCTS 104

DB 1 MSEREGRGKGGKKERGGKPPSAAGSQSPALPPLKEMKSOESAGSKLVLRCTS 60
QY 105 SEYSLRKPWFKNGLNKNKPNQIKQKPKSELINKASLADSGYCMCKVLSKLG 164
DB 61 SEYSLRKPWFKNGLNKNKPNQIKQKPKSELINKASLADSGYCMCKVLSKLG 120
QY 165 DSANANITIVESNEIITGMPASTEGAYVSSSEPIRISYSTEGANTSSSTSTSTTGTSHLV 224
DB 121 DSANANITIVESNEIITGMPASTEGAYVSSSEPIRISYSTEGANTSSSTSTSTTGTSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARCCTENPMKVONOEKAEELYQ 284
DB 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKQCPGFTGARCCTENPMKVONOEKAEELYQ 237
QY 285 KRVLTITGICIAALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNMIANGPHHPN 344
DB 238 KRVLTITGICIAALLVVGIMCVVAYCKTKKQKKLHDLRQSLRSERNMMNMIANGPHHPN 297
QY 345 PPNENVOLVQVSVKNVSSSEHIVERAEATSESTSHYTSHTAHSTVTTQTPSHSWSNGHT 404
DB 298 PPNENVOLVQVSVKNVSSSEHIVERAEATSESTSHYTSHTAHSTVTTQTPSHSWSNGHT 357
QY 405 ESILSESHSVIYVSSVENSRRHSSPTGGPRGLNGTGGPRECNFSLRHARETSDSYRDSPH 464
DB 358 ESILSESHSVIYVSSVENSRRHSSPTGGPRGLNGTGGPRECNFSLRHARETSDSYRDSPH 417
QY 465 SERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAYSPFMEERPLLL 524
DB 418 SERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPPVSSMTVSMPSMAYSPFMEERPLLL 477
QY 525 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETQVEYPAQEPVK 584
DB 478 VTPPLREKKFDHHPQOFSFHHNPAHDSNLSPLASPLRIVEDEYETQVEYPAQEPVK 537
QY 585 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDETVGDTPLFIGNPLAASLE 644
DB 538 LANSRRAKRTKPNGHIANRLEVDNSTSSQSSSESETEDETVGDTPLFIGNPLAASLE 597
QY 645 ATPAFRLADSRTPNAGRFSTQEEIQ 669
DB 598 ATPAFRLADSRTPNAGRFSTQEEIQ 622
RESULT 3
B43273
heregulin, splice form beta 1 - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Aug-2002
C:Accession: B43273; 138406
R:Holmes, W.E.; Sliwkowski, M.X.; Akita, R.W.; Henzel, W.J.; Lee, J.; Park, J.W.; Yan
Science 256, 1205-1210, 1992
A:Title: Identification of heregulin, a specific activator of p185(erbB2).
A:Reference number: A43273; MUID:92271253; PMID:1350381
A:Accession: B43273
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 1-645 <HOL>
R:Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: 138406
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 'A', 95-418, 'F', 420-645 <RES>
A:Cross-references: EMBL:U02328; NID:g408406; PIDN:AAA19953.1; PID:g408407
C:Genetics:
A:Gene: GDB:HGL
A:Cross-references: GDB:132656; OMIM:142445
A:Map position: 8p22-8p11
C:Superfamily: human heregulin; EGF homology
C:Keywords: alternative splicing

RESULT 5

161718
neu differentiation factor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 02-Aug-2002
C:Accession: I61718; I61721; I61720
R:Men, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.;
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation fa
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I61718
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-636 <RES>
A:Cross-references: EMBL:U02318; NID:g408386; PIDN:AAA19943.1; PID:g408387
A:Accession: I61721
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-444, 'A', 446-636 <RES>
A:Cross-references: EMBL:U02321; NID:g408392; PIDN:AAA19946.1; PID:g408393
A:Accession: I61720
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-298, 386, 'V', 388, 'TR', 391 <RES>
A:Cross-references: EMBL:U02320; NID:g408390; PIDN:AAA19945.1; PID:g408391
C:Superfamily: human heregulin; EGF homology
F:182-221/Domain: EGF homology <EGF>

Query Match 84.2%; Score 2922; DB 2; Length 636;
Best Local Similarity 90.4%; Pred. No. 1.7e-155;
Matches 565; Conservative 23; Mismatches 33; Indels 4; Gaps 2;
QY 45 MSERKEGRGKGKGGKRGSGKPPSAAGSQSPALPRLKEMKKSQSAAGSKLVLRCEYS 104
Db 1 MSERKEGRGKGKGGKRGSGKPPSAAGSQSPALPRLKEMKKSQSAAGSKLVLRCEYS 60
QY 105 SEYSSLRFWKFWNGNENLNKPNQIKQKPKGSELINKASLADSGEYMKVSKLGN 164
Db 61 SEYSSLRFWKFWNGNENLNKPNQIKQKPKGSELINKASLADSGEYMKVSKLGN 120
QY 165 DSANANITVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTGTGSHLV 224
Db 121 DSANANITVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTGTGSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKCPGFTGARTENVMKYNQKAEELYQ 284
Db 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKCPGFTGARTENVMKYNQKAEELYQ 237
285 KRLVTITGICIAALLVVGIMCVAYCKTKKQKRLDRLQSLRSERNMMNTANGPHPN 344
Db 238 KRLVTITGICIAALLVVGIMCVAYCKTKKQKRLDRLQSLRSERNMMNTANGPHPN 297
QY 345 PPNVQVLYNQYVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSNNGHT 404
Db 298 PPNVQVLYNQYVSKNVISSSEHIVERAEATSFSTSHYTSTAHHSTVTQTPSHSNNGHT 357
QY 405 ESILSSSHSVIWMSSVENSRRHSPGPGRLNGTGGPRECSNFURHARETDPDSYRDSH 464
Db 358 ESIVISSESVIWMSSVENSRRHSPGPGRLNGTGGPRECSNFURHARETDPDSYRDSH 416
QY 465 SERIVSAMTTPARMSVPDFTPTSSPKSPPESEMPVSSMTVSNMPSNAVSPFMEERPLLL 524
Db 417 SERIVSAMTTPARMSVPDFTPTSSPKSPPESEMPVSSMTVSNMPSNAVSPFMEERPLLL 476
QY 525 VTPRLREKKFDHPHQFSSFHNPADHNSLSPASPLRIVEDEEYTTQYEYPAQSPVKK 584
Db 477 VTPRLREKKYDHPHQFSSFHNPADHNSLSPASPLRIVEDEEYTTQYEYPAQSPVKK 536
QY 585 LANSRAKTKPNGHIANLEVDNSTSSSSSESETEDEYVGDTPFGLIGNPLAASLE 644
Db 537 VTNSRAKTKPNGHIANLEVDNSTSSSSSESETEDEYVGDTPFGLIGNPLAASLE 596
QY 645 ATPAFRLAERTNPAGRFSTQEELQ 669

RESULT 6

161722
neu differentiation factor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 02-Aug-2002
C:Accession: I61722
R:Men, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.
Mol. Cell. Biol. 14, 1909-1919, 1994
A:Title: Structural and functional aspects of the multiplicity of Neu differentiation
A:Reference number: A56210; MUID:94158863; PMID:7509448
A:Accession: I61722
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-662 <RES>
A:Cross-references: EMBL:U02322; NID:g408394; PIDN:AAA19947.1; PID:g408395
C:Superfamily: human heregulin; EGF homology
F:182-221/Domain: EGF homology <EGF>

Query Match 84.1%; Score 2917; DB 2; Length 662;
Best Local Similarity 87.3%; Pred. No. 3.5e-155;
Matches 566; Conservative 25; Mismatches 33; Indels 24; Gaps 3;
QY 45 MSERKEGRGKGKGGKRGSGKPPSAAGSQSPALPRLKEMKKSQSAAGSKLVLRCEYS 104
Db 1 MSERKEGRGKGKGGKRGSGKPPSAAGSQSPALPRLKEMKKSQSAAGSKLVLRCEYS 60
QY 105 SEYSSLRFWKFWNGNENLNKPNQIKQKPKGSELINKASLADSGEYMKVSKLGN 164
Db 61 SEYSSLRFWKFWNGNENLNKPNQIKQKPKGSELINKASLADSGEYMKVSKLGN 120
QY 165 DSANANITVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTGTGSHLV 224
Db 121 DSANANITVESNEITGMPASTEGAYVSESPRISVSTEGANTSSSTSTGTGSHLV 180
QY 225 KCAEKEKTCVNGGECFVMDLSNPSRYLCKCPGFTGARTENVMKYNQKAEELYQ 289
Db 181 KCAEKEKTCVNGGECFVMDLSNPSRYLCKCPGFTGARTENVMKYNQKAEELYQ 240
QY 270 --PMK-----VONQKAEELYQKRVLTITGICIAALLVVGIMCVAYCKTKKQKRLDHR 321
Db 241 EKPLERKLDHSLVKESKAEELYQKRVLTITGICIAALLVVGIMCVAYCKTKKQKRLDHR 300
QY 322 LRQSLRSERNMMNTANGPHPNPPPPENLVQVYVSKNVISSSEHIVERAEATSFSTSHY 381
Db 301 LRQSLRSERNMMNTANGPHPNPPPPENLVQVYVSKNVISSSEHIVERAEATSFSTSHY 360
QY 382 TSTAHHSTVTQTPSHSNNGHTESILSSSHSVIWMSSVENSRRHSPGPGRLNGTGG 441
Db 361 TSTAHHSTVTQTPSHSNNGHTESILSSSHSVIWMSSVENSRRHSPGPGRLNGTGG 420
QY 442 PRECSNFURHARETDPDSYRDSHSPSERIVSAMTTPARMSVPDFTPTSSPKSPPESEMPVPS 501
Db 421 PRD-NSFLRHARETDPDSYRDSHSPSERIVSAMTTPARMSVPDFTPTSSPKSPPESEMPVPS 479
QY 502 SMTVSNMPSNAVSPFMEERPLLLVTPRLREKKFDHPHQFSSFHNPADHNSLSPASPL 561
Db 480 SMTVSNMPSNAVSPFMEERPLLLVTPRLREKKYDHPHQFSSFHNPADHNSLSPASPL 539
QY 562 RIVEDEEYTTQYEYPAQSPVKKLANRRAKTKPNGHIANLEVDNSTSSSSSESETE 621
Db 540 RIVEDEEYTTQYEYPAQSPVKKLANRRAKTKPNGHIANLEVDNSTSSSSSESETE 599
QY 622 EDERVGEDTPFGLIGNPLAASLEAVAPAFRLAERTNPAGRFSTQEELQ 669
Db 600 EDERVGEDTPFGLIGNPLAASLEAVAPAFRLAERTNPAGRFSTQEELQ 647

RESULT 7

A45769

A: Experimental source: PC-12 cell
C: Comment: This protein is a member of the epidermal growth factor family. It is functioning in the differentiation of MDA-MB-453 cells.
C: Superfamily: human ErbB kinase activator alpha, brain and thymus; EGF homology F;361-397/Domain: EGF homology <EGF>

Query Match 28.9%; Score 1001.5; DB 2; Length 868;
Best Local Similarity 36.8%; Pred. No. 2.6e-48;
Matches 248; Conservative 102; Mismatches 184; Indels 139; Gaps 21;
QY 55 GKKKKRGSGKKPESAGSQSPALPPRLKEMKQESAGSKLVLCETSSSEYSSLRFKW 114
Db 234 GNKIKVEKILCTDCA-----RPLKKMKASQGEVGEKQSLKCEAAGNPQPSYRW 286
QY 115 FRNGNELNRNKPQNIQKPKG--SELINKASLADSGEYMKVSKLGNDSASANIT 172
Db 287 FKDGKELNRS--RDRIKYGNGRKNRLQFNKRVEDAGEYVCEAEINLGRDVTGRHL 343
QY 173 IVESNELITGMPASTGAYVSSSESPRISVSTEGANTSSSTSTTGTSHLVKCAEKEKT 232
344 V-----NSVSTTLSSWSG--HARRCNETAKS 367
QY 233 FCVNGCEFMVKDLSNPSRYLCKCPGFTGARTENVPKQV-----NOEKAEE 281
Db 368 YCVNGGVCIYIEGINOLS---CKCPNGFFGQRCLEKPLRLYMPDPKQKHLGFELKEAE 424
QY 282 LYOKRVLTTIGICVALLVGVIMCVWAYCKTKQKKLHDRLQSLRSENNMMNIANGPH 341
Db 425 LYOKRVLTTIGICVALLVGVIMCVWAYCKTKQRRQMHHLRQNMCPAHQ--RSLANGPS 483
QY 342 HNPPEPVQVQVYKKNVISEHIVERAEETSFSTSHYSTAHSSTVQT-----PS 396
Db 484 HPLDPEEIOMAD-YISKVNPATDVRIRAEATTFSGSHSCSPSHHCSTATPTSSHRHES 542
QY 397 HWSNGHTESIASESHSVIWMSEVNSRHSSPT--GGPRGLNGTGGPRECNGLRHARET 455
Db 543 HTWSLERSSELSDSOGIMLSVGTCKNSCPACVEARARAAAYSQE-----RRRAAMP 598
QY 456 P-----DSVRDSPHSRYVSAMTPARMSPVDFHTPSPKSPSEMSPPVSSMTVSPSM 510
Db 599 PYHDSIDSLRSDSPHSRYVSALTTPARLSPVDFHYSLATQVPTFEITSPNSAHAVSLPPA 658
QY 511 A-VSPFMEERERPLLVTPLREKEDHP-----QOFSSFHENPAHDS--- 553
Db 659 APISYRLAEQQPLL-----RHPAPPGPGPGADMQRSYDSYYPAAAGPGPR 705
QY 554 -----NSLPASPLRIVEDEYETQYEPAQEPVKKL-----ANSRRAKRTK 595
Db 706 RGACALGSGSLSPASPFRIPEDDEYETQECAPPPPPRPRTRGASRRTSAGPRWRRSR 765
QY 596 PNGHTIANR-----LEVDNTSSOSSNSESETEDEVRG-----EDTPFLGIQNPL-AASL 643
Db 766 LNGLAAQARAARDLSLSGSGCGSASASDDADDGALAAESTPFLGLRAAHDLRS 825
QY 644 EATPAFLADSR 656
Db 826 DSPPLCGPAADSR 838

RESULT 15

A56210
neu differentiation factor - rat (fragment)
C: Species: Rattus norvegicus (Norway rat)
C: Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 02-Aug-2002
C: Accession: A56210
R: Wen, D.; Suggs, S.V.; Karunakaran, D.; Liu, N.; Cupples, R.L.; Luo, Y.; Janssen, A.M.; Mol. Cell. Biol. 14, 1509-1919, 1994
A: Title: Structural and functional aspects of the multiplicity of Neu differentiation factor
A: Reference number: A56210; MUID: 94158863; PMID: 7509448
A: Accession: A56210
A: Status: preliminary;
A: Molecule type: mRNA
A: Residues: 1-230 <RES>

A: Cross-references: EMBL:U02315; NID:g408380; PIDN:AAA19940.1; PID:g408381
C: Superfamily: human heregulin; EGF homology

Query Match 28.4%; Score 985; DB 2; Length 230;
Best Local Similarity 90.7%; Pred. No. 4.1e-48;
Matches 194; Conservative 3; Mismatches 17; Indels 0; Gaps 0;
QY 56 GKKKKRGSGKKPESAGSQSPALPPRLKEMKQESAGSKLVLCETSSSEYSSLRFKW 115
Db 1 KGKKKDRGSRGKPGPAEGDPSPALPPRLKEMKQESAGSKLVLCETSSSEYSSLRFKW 60
QY 116 KGNELNRNKPQNIQKPKGSELINKASLADSGEYMKVSKLGNDSASANIT 175
Db 61 KGNELNRNKPQNIQKPKGSELINKASLADSGEYMKVSKLGNDSASANIT 120
QY 176 SNEITGMPASTGAYVSSSESPRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTECV 235
Db 121 SNEITGMPASTGAYVSSSESPRISVSTEGANTSSSTSTTGTSHLVKCAEKEKTECV 180
QY 236 NGGECFMVKDLSNPSRYLCKCPGFTGARTENVP 269
Db 181 NGGECFTVKDLSNPSRYLCKCPNEFTGDRCONV 214

Search completed: July 9, 2003, 10:48:08
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:23:15 ; Search time 15 Seconds
(without alignments)
1849.844 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470

Sequence: 1 ARAPQRGRSLSPSRDKLPN.....RLASRTNPAGRFSTQEEIQ 669

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3242	93.4	639	1 NRGL_HUMAN	Q02297 h pro-neure
2	2917	84.1	662	1 NRGL_RAT	P43322 r pro-neure
3	2314	66.7	602	1 NRGL_CHICK	Q05199 gallus gall
4	2086.5	60.1	677	1 NRGL_XENLA	O93383 xenopus lae
5	1013.5	29.2	850	1 NRGL_HUMAN	O14511 homo sapien
6	1001.5	28.9	868	1 NRGL_RAT	O35569 rattus norv
7	987.5	28.5	756	1 NRGL_MOUSE	P56974 mus musculu
8	301	8.7	713	1 NRGL_MOUSE	O35181 mus musculu
9	300	8.6	720	1 NRGL_HUMAN	P56975 homo sapien
10	273	7.9	296	1 SHDF_HUMAN	Q15491 homo sapien
11	174.5	5.0	704	1 NP14_RAT	P41777 rattus norv
12	165	4.8	1091	1 NCAL_CHICK	P13590 gallus gall
13	162	4.7	1306	1 MSB2_YEAST	P32334 saccharomyc
14	160.5	4.6	699	1 NP14_HUMAN	Q14978 homo sapien
15	157.5	4.5	3924	1 ANK2_HUMAN	Q01484 homo sapien
16	156	4.5	992	1 FRQ_LEPAU	Q01115 leptosphaer
17	154	4.4	1943	1 PC15_MOUSE	Q993p1 mus musculu
18	152	4.4	2843	1 APC_HUMAN	P25054 homo sapien
19	151.5	4.4	785	1 IF16_HUMAN	Q16666 homo sapien
20	150	4.3	1189	1 YJH6_YEAST	P47035 saccharomyc
21	149	4.3	906	1 CENC_MOUSE	P49452 mus musculu
22	148	4.3	633	1 MLH_TETTH	P40631 tetrahymena
23	147.5	4.3	1794	1 YAV1_SCHPO	Q10172 schizosacch
24	147	4.2	862	1 CD22_MOUSE	P35329 mus musculu
25	146.5	4.2	1092	1 NCAL_XENLA	P36335 xenopus lae
26	146	4.2	1636	1 BUD3_YEAST	P25558 saccharomyc
27	145.5	4.2	1115	1 NCAL_MOUSE	P35395 mus musculu
28	145	4.2	1411	1 TCOP_HUMAN	Q13428 homo sapien
29	143	4.1	4393	1 PGBM_HUMAN	P98160 homo sapien
30	142	4.1	3329	1 BRC2_MOUSE	P97929 mus musculu
31	140	4.0	1849	1 IG4A_HAEIN	P45386 haemophilus
32	139.5	4.0	323	1 FCGC_HUMAN	P31995 homo sapien
33	139.5	4.0	601	1 YFK5_SCHPO	P87132 schizosacch

RESULT 1

NRGL_HUMAN

ID NRGL_HUMAN

AC Q02297; Q02298; Q02299; Q12779; Q12780; Q12781; Q12782; Q12783;

AC Q12784; Q07110; Q07111; Q9UPE3; Q14667;

DT 16-OCT-2001 (Rel. 40, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Pro-neuregulin-1 precursor (Pro-NRGL) [Contains: Neuregulin-1 (Neu

DE differentiation factor) (Heregulin) (HRG) (Breast cancer cell

DE differentiation factor p45) (Acetylcholine receptor inducing activity)

DE (ARIA) (Sensory and motor neuron-derived factor) (Glial growth

DE factor)]

GN NRGL OR HGL OR NDF OR HRGA OR GGF OR SMDF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA1; BETA2 AND BETA3), AND

RP PARTIAL SEQUENCE.

RX MEDLINE=92271253; PubMed=1350381;

RA Holmes W.E., Sliwowski M.X., Akita R.W., Henzel W.J., Lee J.,

RA Papp W.W., Yansura D., Abadi N., Raab H., Lewis G.D., Shepard H.M.,

RA Kuang W.-J., Wood W.I., Goeddel D.V., Vanden R.L.;

RL Identification of heregulin, a specific activator of p185erbB2.;

RN Science 256:1205-1210(1992).

RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; ALPHA2B; ALPHA3; BETA1; BETA2

RP AND BETA3).

RC TISSUE=pituitary, and kidney adenocarcinoma;

RC MEDLINE=94158863; PubMed=7509448;

RA Wen D., Suggs S.V., Karunakaran D., Liu N., Cupples R.L., Luo Y.,

RA Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,

RA Meng S.-Y., Lu H.S., Hu S., Chang D., Yang W., Yanigahara D.,

RA Koski R.A., Yarden Y.;

RT "Structural and functional aspects of the multiplicity of Neu

RT differentiation factors.;"

RL Mol. Cell. Biol. 14:1909-1919(1994).

RN [3]

RP SEQUENCE FROM N.A. (ISOFORM ALPHA).

RX MEDLINE=92208945; PubMed=1348215;

RA Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,

RA Levy R.B., Yarden Y.;

RT "Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein

RT that induces differentiation of mammary tumor cells.;"

RL Cell 69:205-216(1992).

RN [4]

RP SEQUENCE FROM N.A. (ISOFORMS BETA3 AND GGF2).

RC TISSUE=Brain;

RC MEDLINE=93205115; PubMed=8096067;

RA Marchionni M.A., Gooden A.D.J., Chen M.S., Birmingham-McDonogh O.,

RA Kirk C., Hendricks M., Danahy F., Misumi D., Sudhalter J.,

RA Kobayashi K., Wrblewski D., Lynch C., Baldasare M., Hiles I.,

RA Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,

RA Waterfield M.D., Stroobant P., Gwynne D.;

ALIGNMENTS

34	139.5	4.0	1133	1 EGF_RAT	P07522 rattus norv
35	139.5	4.0	2717	1 ZEP1_HUMAN	P15822 homo sapien
36	139	4.0	862	1 SMP2_YEAST	P32567 saccharomyc
37	138.5	4.0	2845	1 APC_MOUSE	Q61315 mus musculu
38	138	4.0	764	1 ICCR_DROME	Q08180 drosophila
39	137	3.9	1468	1 FMN1_MOUSE	Q05860 mus musculu
40	136.5	3.9	1906	1 KMLS_CHICK	P11799 gallus gall
41	136.5	3.9	2842	1 APC_RAT	P70478 rattus norv
42	135.5	3.9	936	1 FHL1_YEAST	P39521 saccharomyc
43	135.5	3.9	952	1 YK15_CAEEL	P46012 caenorhabdi
44	135.5	3.9	1087	1 POM1_SCHPO	Q09690 schizosacch
45	135	3.9	1163	1 RTN4_RAT	Q91k11 rattus norv

RT *Glial growth factors are alternatively spliced erbB2 ligands
 RL expressed in the nervous system.*;
 RL Nature 362:312-318(1993).
 [5]
 RP SEQUENCE FROM N.A. OF GAMMA-HEREGULIN FUSION PROTEIN.
 RC TISSUE-Breast cancer;
 RX MEDLINE-97472144; PubMed-9333014;
 RA Schafer G., Fitzpatrick V.D., Sliwkowski M.X.;
 RT "Gamma-hergulin: a novel heregulin isoform that is an autocrine
 RL growth factor for the human breast cancer cell line, MDA-MB-175";
 RL Oncogene 15:1385-1394(1997).
 [6]
 RP SEQUENCE OF 1-210 FROM N.A.
 RA Schoumacher F., Herzer S., Flury N., Kueng W., Mueller H.,
 RA Eppenberger U.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 [7]
 RP SEQUENCE OF 19-27.
 RX MEDLINE-93366731; PubMed-7689552;
 RA Culouscou J.-M., Plozman G.D., Carlton G.W., Green J.M., Shoyab M.;
 RT "Characterization of a breast cancer cell differentiation factor that
 RL specifically activates the HER4/p180erbB4 receptor.";
 RL J. Biol. Chem. 268:18407-18410(1993).
 [8]
 RP CHROMOSOMAL TRANSLOCATION
 RX MEDLINE-99455251; PubMed-10523851;
 RA Wang X.-Z., Jolicoeur E.M., Conte N., Chaffanet M., Zhang Y.,
 RA Mozziconacci M.-J., Feiner H., Birnbaum D., Pebusque M.-J., Ron D.;
 RT "Gamma-hergulin is the product of a chromosomal translocation fusing
 RL the DOC4 and HGL/NRG1 genes in the MDA-MB-175 breast cancer cell
 RL line.";
 RL Oncogene 18:5718-5721(1999).
 [9]
 RP CHROMOSOMAL TRANSLOCATION.
 RX MEDLINE-20065180; PubMed-10597312;
 RA Liu X., Baker E., Eyre H.J., Sutherland G.R., Zhou M.;
 RT "Gamma-hergulin: a fusion gene of DOC-4 and neuregulin-1 derived from
 RL a chromosome translocation.";
 RL Oncogene 18:7110-7114(1999).
 [10]
 RP STRUCTURE BY NMR OF 175-241 (ISOFORM ALPHA).
 RX MEDLINE-94341264; PubMed-8062828;
 RA Nagata K., Kohda D., Hatanaka H., Ichikawa S., Matsuda S.,
 RA Yamamoto T., Suzuki A., Inagaki F.;
 RT "Solution structure of the epidermal growth factor-like domain of
 RL heregulin-alpha, a ligand for p180erbB-4.";
 RL EMBO J. 13:3517-3523(1994).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM
 CC DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF
 CC EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING
 CC EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING
 CC THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING
 CC LOBOLOALVEOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND
 CC AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING
 CC SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE
 CC MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART.
 CC -1- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
 CC REGION OF LIMK1 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS
 CC A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE. THE SECRETED
 CC ISOFORM GGF2 HAS A SIGNAL PEPTIDE. THE ISOFORM BETA3 MAY BE
 CC NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 10 ISOFORMS: ALPHA (SHOWN HERE);
 CC ALPHAIIA, ALPHAIIB, ALPHAIIC, BETA1, BETA2, BETA3/GGFHFB1,
 CC GGF2/GGFHFP2 AND SMDF (AC Q15491); ARE PRODUCED BY ALTERNATIVE
 CC SPLICING. THEY HAVE BEEN CLASSIFIED AS TYPE I NRGS (VARIANTS WITH
 CC AN IG DOMAIN AND A GLYCOSYLATION DOMAIN; ALPHA AND BETA), TYPE II
 CC NRGS (VARIANTS WITH AN IG DOMAIN BUT NO GLYCOSYLATION DOMAIN;
 CC GGF2) AND TYPE III NRGS (VARIANTS WITH A CYS-RICH DOMAIN; SMDF).
 CC

CC ALL THESE ISOFORMS PERFORM DISTINCT TISSUE-SPECIFIC FUNCTIONS.
 CC -1- TISSUE SPECIFICITY: TYPE I ISOFORMS ARE THE PREDOMINANT FORMS
 CC EXPRESSED IN THE ENDOCARDIUM. ISOFORM ALPHA IS EXPRESSED IN
 CC BREAST, OVARY, TESTIS, PROSTATE, HEART, SKELETAL MUSCLE, LUNG,
 CC PLACENTA LIVER, KIDNEY, SALIVARY GLAND, SMALL INTESTINE AND BRAIN,
 CC BUT NOT IN UTERUS, STOMACH, PANCREAS, AND SPLEEN. ALPHAIIB IS THE
 CC PREDOMINANT FORM IN MESENCHYMAL CELLS AND IN NONNEURONAL ORGANS,
 CC WHEREAS BETA1 IS THE MAJOR NEURONAL FORM. BETA3 IS EXPRESSED IN
 CC SPINAL CORD AND BRAIN. GGF2 IS THE MAJOR FORM IN SKELETAL MUSCLE
 CC CELLS; IN THE NERVOUS SYSTEM IT IS EXPRESSED IN SPINAL CORD AND
 CC BRAIN. ALSO DETECTED IN ADULT HEART, PLACENTA, LUNG, LIVER,
 CC KIDNEY, AND PANCREAS.
 CC -1- DEVELOPMENTAL STAGE: DETECTABLE AT EARLY EMBRYONIC AGES.
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN.
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM.
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- DISEASE: INVOLVED IN A RARE T(8;11) CHROMOSOMAL TRANSLOCATION THAT
 CC FUSES THE 5' END OF OD24 TO NRG1 (ISOFORM BETA3). THE PRODUCT OF
 CC THIS TRANSLOCATION WAS FIRST THOUGHT TO BE AN ALTERNATIVELY
 CC SPLICED ISOFORM, CALLED GAMMA-HEREGULIN. GAMMA-HEREGULIN IS A
 CC SOLUBLE ACTIVATING LIGAND FOR THE ERBB2-ERBB3 RECEPTOR COMPLEX AND
 CC ACTS AS AN AUTOCRINE GROWTH FACTOR IN A SPECIFIC BREAST CANCER
 CC CELL LINE (MDA-MB-175). NOT DETECTED IN BREAST CARCINOMA SAMPLES,
 CC INCLUDING DUCTAL, LOBULAR, MEDULLARY, AND MUCINOUS HISTOLOGICAL
 CC TYPES, NEITHER IN OTHER BREAST CANCER CELL LINES.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M94165; AAA58638.1; -
 CC EMBL; M94166; AAA58639.1; -
 CC EMBL; M94167; AAA58640.1; -
 CC EMBL; M94168; AAA58641.1; -
 CC EMBL; L12261; AAB59338.1; -
 CC EMBL; U02325; AAA19950.1; -
 CC EMBL; U02326; AAA19951.1; -
 CC EMBL; U02327; AAA19952.1; -
 CC EMBL; U02328; AAA19953.1; -
 CC EMBL; U02329; AAA19954.1; -
 CC EMBL; U02330; AAA19955.1; -
 CC EMBL; L12260; AAB59622.1; -
 CC EMBL; AF026146; RAD01795.1; -
 CC EMBL; AF009227; AAC51756.1; ALT_INIT.
 CC PDB; 1HRE; 15-OCT-94.
 CC PDB; 1HRF; 15-OCT-94.
 CC Genew; HGNC:7997; NRG1.
 CC MIM; 142445; -
 CC InterPro; IPR003006; EGF-like.
 CC InterPro; IPR002154; Neuregulin.
 CC Pfam; PF00008; EGF; 1.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF02158; Neuregulin; 1.
 CC PRINTS; PR01089; NEUREGULIN.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; 1.
 CC Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
 CC

KW

Query Match 93.4%; Score 3242; DB 1; Length 639;
Best Local Similarity 100.0%; Pred. No. 5.6e-173;
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 SERKEGRGKGGKKGKSPESAGSOSQSPALPRKEMKQSAAGSKVLVRCETSS 105
DB 1 SERKEGRGKGGKKGKSPESAGSOSQSPALPRKEMKQSAAGSKVLVRCETSS 60
QY 106 EYSSLRKFWKNGNKLNRKNPQNIQKPKCKSELINKASLADSGEYMKVSKLGN 165
DB 61 EYSSLRKFWKNGNKLNRKNPQNIQKPKCKSELINKASLADSGEYMKVSKLGN 120
QY 166 SASANITIVESNEIITGMPASTEGAYVSESPISRVSTEGANTSSSTSTGTS 225
DB 121 SASANITIVESNEIITGMPASTEGAYVSESPISRVSTEGANTSSSTSTGTS 180
QY 226 CAEKERTCVNGGECFMVKDLSNPSRYLCKPGFTGARTENPMKVONQKAEELV 285
DB 181 CAEKERTCVNGGECFMVKDLSNPSRYLCKPGFTGARTENPMKVONQKAEELV 240
QY 286 RVLITIGICIAALLVVGIMCVAYCKTKQKRLHDLRQSLRSERNMMNTANGPH 345
DB 241 RVLITIGICIAALLVVGIMCVAYCKTKQKRLHDLRQSLRSERNMMNTANGPH 300
QY 346 PENVOLNQYVSKNVISSHIVERAETSFTSHYTSTAHHSTVTQTPSHSNGHT 405
DB 301 PENVOLNQYVSKNVISSHIVERAETSFTSHYTSTAHHSTVTQTPSHSNGHT 360
QY 406 SILSESHSVIVMSSVENRHSPTGPRGLNGTGGPRECSFLRHARETDPYSD 465
DB 361 SILSESHSVIVMSSVENRHSPTGPRGLNGTGGPRECSFLRHARETDPYSD 420
QY 466 ERYVSAMTTPARMSVPDFTSSPKSPSEMSPPVSSMTVSNPMAVSPFMEER 525
DB 421 ERYVSAMTTPARMSVPDFTSSPKSPSEMSPPVSSMTVSNPMAVSPFMEER 480
QY 526 TPRRLREKFDHPQOFSFHNPADNSLSPASPLRIVEDEYETQYEYPAQBPVK 585
DB 481 TPRRLREKFDHPQOFSFHNPADNSLSPASPLRIVEDEYETQYEYPAQBPVK 540
QY 586 ANSRRAKTKPNHGLANLEVDNNTSSSNSESETEDEYETQYEYPAQBPVK 645
DB 541 ANSRRAKTKPNHGLANLEVDNNTSSSNSESETEDEYETQYEYPAQBPVK 600
QY 646 TPAFLADSRTPNAGRFSTQEEIQ 669
DB 601 TPAFLADSRTPNAGRFSTQEEIQ 624

RESULT 2

NRGI_RAT STANDARD; PRT: 662 AA.
AC P43322; P43323; P43324; P43325; P43326; P43327; P43328;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pro-neuregulin-1 precursor (Pro-NRG1) [Contains: Neuregulin-1 (Neu
differentiation factor) (Heregulin) (HRG) (Acetylcholine receptor
inducing activity) (ARIA) (Sensory and motor neuron-derived factor)
(Glial growth factor)].
GN NRGI OR NDF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE-Fibroblast;
RX MEDLINE=94158863; PubMed=7509448;
RA Wen D., Suggs S.V., Karunakaran D., Liu N., Cupples R.L., Luo Y.,
Janssen A.M., Ben-Baruch N., Trollinger D.B., Jacobsen V.L.,
Meng S.-Y., Lu H.S., Hu S., Chang D., Yang D., Yanigahara D.,

Koski R.A., Yarden Y.;
"Structural and functional aspects of the multiplicity of Neu
differentiation factors";
Mol. Cell. Biol. 14:1909-1919(1994).
[2]
SEQUENCE FROM N.A. (ISOFORM ALPHA2C/NDFA4), AND PARTIAL SEQUENCE.
RP TISSUE-Fibroblast;
RX MEDLINE=92257596; PubMed=1349853;
RA Wen D., Peles E., Cupples R., Suggs S.V., Bacus S.S., Luo Y.,
Trail G., Hu S., Silbiger S.M., Levy R.B., Koski R.A., Lu H.S.,
Yarden Y.;
"A new differentiation factor: a transmembrane glycoprotein containing
an EGF domain and an immunoglobulin homology unit";
Cell 69:559-572(1992).
[3]
SEQUENCE OF 14-36.
RP MEDLINE=92208945; PubMed=1348215;
RA Peles E., Bacus S.S., Koski R.A., Lu H.S., Wen D., Ogden S.G.,
Levy R.B., Yarden Y.;
"Isolation of the neu/HER-2 stimulatory ligand: a 44 kd glycoprotein
that induces differentiation of mammary tumor cells";
Cell 69:205-216(1992).
[4]
REGULATION OF PROCESSING (ISOFORM ALPHA2C/NDFA4).
RP MEDLINE=99069430; PubMed=9852099;
RA Liu X., Hwang H., Cao L., Wen D., Liu N., Graham R.M., Zhou M.;
"Release of the neuregulin functional polypeptide requires its
cytoplasmic tail";
J. Biol. Chem. 273:34335-34340(1998).
[5]
INTERACTION WITH LIMK1.
RP MEDLINE=98352096; PubMed=9685409;
RA Wang J.Y., Frenzel K.E., Wen D., Falls D.L.;
"Transmembrane neuregulins interact with LIM kinase 1, a cytoplasmic
protein kinase implicated in development of visuospatial cognition";
J. Biol. Chem. 273:20525-20534(1998).
CC -!- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
ACTIVATION OF THE ERBB RECEPTORS. THE MULTIPLE ISOFORMS PERFORM
DIVERSE FUNCTIONS SUCH AS INDUCING GROWTH AND DIFFERENTIATION OF
EPITHELIAL, GLIAL, NEURONAL, AND SKELETAL MUSCLE CELLS; INDUCING
EXPRESSION OF ACETYLCHOLINE RECEPTOR IN SYNAPTIC VESICLES DURING
THE FORMATION OF THE NEUROMUSCULAR JUNCTION; STIMULATING
LOBULOVOLAR BUDDING AND MILK PRODUCTION IN THE MAMMARY GLAND
AND INDUCING DIFFERENTIATION OF MAMMARY TUMOR CELLS; STIMULATING
SCHWANN CELL PROLIFERATION; IMPLICATION IN THE DEVELOPMENT OF THE
MYOCARDIUM SUCH AS TRABECULATION OF THE DEVELOPING HEART (BY
SIMILARITY).
CC -!- SUBUNIT: THE CYTOPLASMIC DOMAIN INTERACTS WITH THE LIM DOMAIN
REGION OF LIMK1.
CC -!- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
BOUND FORM DOES NOT SEEM TO BE ACTIVE.
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; ALPHA2A/NDFA38,
ALPHA2B/NDFA19, ALPHA2C/NDFA4, BETA1, BETA2/NDFA40, BETA2A/NDFA22,
BETA3/NDFA4 AND BETA4/NDFA42A (SHOWN HERE); ARE PRODUCED BY
ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. MOST TISSUES CONTAIN ALPHA2A
AND ALPHA2B ISOFORMS. ALPHA2A AND BETA2 ARE THE PREDOMINANT FORMS
IN MESENCHYMAL AND NONNEURONAL ORGANS. BETA1 IS ENRICHED IN BRAIN
AND SPINAL CORD, BUT NOT IN MUSCLE AND HEART. ALPHA2C IS HIGHLY
EXPRESSED IN SPINAL CORD, MODERATELY IN LUNG, BRAIN, OVARY, AND
STOMACH, IN LOW AMOUNTS IN THE KIDNEY, SKIN AND HEART AND NOT
DETECTED IN THE LIVER, SPLEEN, AND PLACENTA.
CC -!- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION OF
TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
DIMERIZATION.
CC -!- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
DOMAIN.
CC -!- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR

FORM.

CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U02315; AAA19940.1; -

DR EMBL; U02316; AAA19941.1; -

DR EMBL; U02317; AAA19942.1; -

DR EMBL; U02318; AAA19943.1; -

DR EMBL; U02319; AAA19944.1; -

DR EMBL; U02320; AAA19945.1; -

DR EMBL; U02321; AAA19946.1; -

DR EMBL; U02322; AAA19947.1; -

DR EMBL; U02323; AAA19948.1; -

DR EMBL; U02324; AAA19949.1; -

DR EMBL; M92430; -; NOT_ANNOTATED_CDS.

DR HSSP; Q12784; IHRE.

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR002154; Neuregulin.

DR Pfam; PF00008; EGF; 1.

DR Pfam; PF00047; Ig; 1.

DR PRINTS; PF01089; Neuregulin; 1.

DR SMART; SM00408; IGC2; 1.

DR SMART; SM00408; IGC2; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; FALSE_NEG.

DR Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;

DR Transmembrane; Multigene family; Alternative splicing.

KW PROPEP 1 13

FT CHAIN 14 662 PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.

FT CHAIN 14 264 NEUREGULIN-1.

FT DOMAIN 14 265 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 266 288 INTERNAL SIGNAL SEQUENCE (POTENTIAL).

FT DOMAIN 289 662 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 50 119 IG-LIKE C2-TYPE DOMAIN.

FT DOMAIN 165 177 SER/THR-RICH.

FT DOMAIN 178 222 EGF-LIKE.

FT DISULFID 57 112 BY SIMILARITY.

FT DISULFID 182 196 BY SIMILARITY.

FT DISULFID 190 210 BY SIMILARITY.

FT DISULFID 212 221 BY SIMILARITY.

FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPPLIC 213 256 PHEFTGDRQNTVYMASFYTSRRKRQETKPLEKRLDHSLV

FT KES -> QPGFGARTENPMKVQIQE (IN ISOFORM

FT ALPHA2A, ISOFORM ALPHA2B AND ISOFORM

FT ALPHA2C).

FT WTSRRKRQETKPLEKRLDHSLVKESK -> KHLGIEPME

FT (IN ISOFORM BETA1).

FT MTSRRKRQETE -> STSTPFLSLPE (IN ISOFORM

FT BETA3).

FT MISSING (IN ISOFORM BETA3).

FT MISSING (IN ISOFORM BETA2 AND ISOFORM

FT BETA2A).

FT PENVQ -> RVTRRG (IN ISOFORM BETA2).

FT MISSING (IN ISOFORM ALPHA2C).

FT YVSAMTTPARMSPVDFHTPSSPKSPSEMSPVSSMTVS

FT -> HNLIAELRNKAYRSCMQIQLSATHLRPSSTHILGFI

FT L (IN ISOFORM ALPHA2B).

FT MISSING (IN ISOFORM ALPHA2B).

FT VARSPPLIC 485 662

FT CONFLICT 90 90 K -> N (IN REF. 2).

FT CONFLICT 137 137 T -> I (IN REF. 2; AA SEQUENCE).

FT CONFLICT 208 208 Y -> S (IN REF. 2).

SQ SEQUENCE 662 AA; 73288 MR; 1C3IABCF2A8EBID5 CRC64;

Query Match 84.1%; Score 2917; DB 1; Length 662;

Best Local Similarity 87.3%; Pred. No. 6.2e-155;

Matches 566; Conservative 25; Mismatches 33; Indels 24; Gaps 3;

QY 45 MSERKEGRKGKGGKKGKGPESAAQSQSPALPRLKEMKQSAAGSKLVLRCEUS 104

DB 1 MSERKEGRKGKGGKKGKGPESAAQSQSPALPRLKEMKQSAAGSKLVLRCEUS 60

QY 105 SEYSSLRFWKFNKNGNELNKNRPQNIKIQKPKGKSELINKASLADSGEYMKVSKLGN 164

DB 61 SEYSSLRFWKFNKNGNELNKNRPQNIKIQKPKGKSELINKASLADSGEYMKVSKLGN 120

QY 165 DSASANITIVESNEITGMPASTEGAYVSSSPIRISVTEGANTSSSTSTGTGSHLV 224

DB 121 DSASANITIVESNEITGMPASTETAYVSSSPIRISVTEGANTSSSTSTGTGSHLI 180

QY 225 KCAERKTEFCVNGGECFVMDLSNPRLCKPCPGFTGARTENV----- 269

DB 181 KCAERKTEFCVNGGECFVMDLSNPRLCKPCPGFTGARTENV----- 240

QY 270 --PMK-----VONQKAEELYQKRVLTITIGICALLVVGIMCVVAYCKTKKQKRLHDR 321

DB 241 EXPLERKLDHSLVKESKAEELYQKRVLTITIGICALLVVGIMCVVAYCKTKKQKRLHDR 300

QY 322 LRQSLASERNMMNTANGPHHPNPENPVOLVQYVSKNVISEHIVERAEFTSFSTSHY 381

DB 301 LRQSLASERNLVNTANGPHHPNPENPVOLVQYVSKNVISEHIVERAEFTSFSTSHY 360

QY 382 TSTAHTSTVTPTPSHWSNGHTESTLSFHSVIVMSSVSHSSPTGPGRLNGTG 441

DB 361 TSTAHTSTVTPTPSHWSNGHTESTLSFHSVIVMSSVSHSSPTGPGRLNGTG 420

QY 442 PRECHNSFLRHARETDPDSRSPHSERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPVVS 501

DB 421 PRD-NSFLRHARETDPDSRSPHSERYVSAMTTPARMSPVDFHTPSSPKSPSEMSPVVS 479

QY 502 SMTVSPMSVAVSPFVEERPLLVTPPRLREKFFDHPHQFSSFHINPAHDSNLSPASPL 561

DB 480 SMTVSPMSVAVSPFVEERPLLVTPPRLREKFFDHPHQFSSFHINPAHDSNLSPASPL 539

QY 562 RIVEDEEYETTOYEYPAQEPVKKLANRRAKTKPNGHIANRLEVDNSTSSSSNSESET 621

DB 540 RIVEDEEYETTOYEYVQEPVKVNSRRAKTKPNGHIANRLEVDNSTSSSSNSESET 599

QY 622 EDVRVEDTPFLGIQNPALASLEATPAFLADSRTPNAGRFSTQEEIQ 669

DB 600 EDVRVEDTPFLGIQNPALASLEAVAPAFRLAERTNPAGRFSTQEEIQ 647

RESULT 3

NRGI_CHICK

ID NRGI_CHICK STANDARD; PRT: 602 AA.

AC Q05199; 073750; 073751; 073752;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Pro-neuregulin-1 precursor (pro-NRGI) [Contains: Neuregulin-1

DE (Acetylcholine receptor inducing activity) (ARIA)].

GN NRGI OR ARIA.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM ARIA), AND PARTIAL SEQUENCE.

RC STRAIN=White leghorn; TISSUE=Brain;

RX MEDLINE=93201602; PubMed=8453670;

RA Falls D.L., Rosen K.M., Corfas G., Lane W.S., Fischbach G.D.;
RT "ARIA, a protein that stimulates acetylcholine receptor synthesis, is
RL a member of the neu ligand family.";
RN Cell 72:801-815(1993).
RP SEQUENCE FROM N.A. (ISOFORMS BETA1A; BETA2A AND BETA2B).
RC TISSUE-BRAIN, and Spinal cord;
RX MEDLINE-98150351; PubMed-9491987;
RA Yang X., Kuo Y., Devay P., Yu C., Role L.;
RT "A cysteine-rich isoform of neuroligin controls the level of
RT expression of neuronal nicotinic receptor channels during
RL synaptogenesis.";
RL Neuron 20:255-270(1998).
CC -1- FUNCTION: DIRECT LIGAND FOR THE ERBB TYROSINE KINASE RECEPTORS.
CC THE MULTIPLE ISOFORMS PERFORM DIVERSE FUNCTIONS: CYSTEIN-RICH
CC DOMAIN CONTAINING ISOFORMS (CRD-NRG) PROBABLY REGULATE THE
CC EXPRESSION OF NICOTINIC ACETYLCHOLINE RECEPTORS AT DEVELOPING
CC INTERNEURONAL SYNAPSES. THE IG-NRG ISOFORM IS REQUIRED FOR THE
CC INITIAL INDUCTION AND/OR MAINTENANCE OF THE MATURE LEVELS OF
CC ACETYLCHOLINE RECEPTORS AT NEUROMUSCULAR SYNAPSES.
CC -1- SUBCELLULAR LOCATION: EXISTS AS A TYPE I MEMBRANE PROTEIN AND AS A
CC PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE MEMBRANE-
CC BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; ARIA/IG-NRG (SHOWN
CC HERE), CRD-NRG-BETA1A, CRD-NRG-BETA2A AND CRD-NRG-BETA2B; ARE
CC PRODUCED BY ALTERNATIVE SPLICING. ARIA CONTAINS AN IG-LIKE DOMAIN,
CC WHEREAS IN THE CRD-NRG (OR NARIA) ISOFORMS, THE EGF-LIKE DOMAIN IS
CC REPLACED BY A CYSTEINE-RICH DOMAIN (CRD).
CC -1- DEVELOPMENTAL STAGE: CRD-NRG ISOFORM IS DETECTED AT EMBRYONIC DAY
CC 4 (ED4) IN BOTH VISCERAL AND SOMATIC MOTOR NEURONS OF SPINAL CORD
CC AND IS HIGHEST AT ED6. IG-NRG ISOFORM IS NOT EXPRESSED UNTIL ED 6
CC IN SPINAL CORD. AT ED 11 BOTH ISOFORMS DISPLAY COMPARABLE LEVELS.
CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
CC DIMERIZATION (BY SIMILARITY).
CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
CC DOMAIN.
CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
CC FORM.
CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L11264; AAA49037.1; -;
CC EMBL; AF045654; AAC05670.1; -;
CC EMBL; AF045655; AAC05671.1; -;
CC EMBL; AF045656; AAC05672.1; -;
CC HSSP; Q12784; 1HRE.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR002154; Neuregulin.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF02158; Neuregulin; 1.
CC PRINTS; PR01089; NEUREGULIN.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00408; Ig_c2; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE_NEG.
CC Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;
CC Transmembrane; Alternative splicing.

FT	CHAIN	1	602	PRO-NEUREGULIN-1, MEMBRANE-BOUND FORM.
FT	CHAIN	1	205	NEUREGULIN-1.
FT	DOMAIN	1	206	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	207	229	INTERNAL SIGNAL SEQUENCE (POTENTIAL).
FT	DOMAIN	230	602	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	42	112	IG-LIKE C2-TYPE DOMAIN.
FT	DOMAIN	125	136	SER/THR-RICH.
FT	DOMAIN	137	181	EGF-LIKE.
FT	DISULFID	49	105	BY SIMILARITY.
FT	DISULFID	141	155	BY SIMILARITY.
FT	DISULFID	149	169	BY SIMILARITY.
FT	DISULFID	171	180	BY SIMILARITY.
FT	CARBOHYD	21	21	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	113	113	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPPLIC	1	127	WNTSEGLQVSLAPOTDQVNSVNTVPKLNKRNORVAV
FT				GOKVLCRTSEYSPALREFKLNKKEITKRNENKIPK
FT				KOKYSELHIYRATLADAGEACRVSSKLGNDSTKASVIT
FT				PHREDSRVGVAGLASTCCVCLAEARLKGCLNSEKICIAPI
FT				LACLLSLCICLAGLKWVDFKIFEDYSDTHLDGRIGODPR
FT				STVDPTLSANWPEVASPPFIQPSLESKAETVQTDSSLV
FT				PSRPTLOPSLXNRILDVGLWSSATPSLSPSLEPTASQAO
FT				ATENIOLAPKLS (IN ISOFORM BETA1A, ISOFORM
FT				BETA2A AND ISOFORM BETA2B).
FT				MISSING (IN ISOFORM BETA2A AND ISOFORM
FT				BETA2B).
FT	VARSPPLIC	191	198	VSAMTTPARMSPVDFHTP -> HTPPTSLLAGKVLRSV
FT	VARSPPLIC	388	405	(IN ISOFORM BETA2B).
FT	VARSPPLIC	406	602	MISSING (IN ISOFORM BETA2B).
FT	SEQUENCE	602 AA;	67453 MW;	41833COE56CED5346 CRC64;
QY	Query Match		66.7%;	Score 2314; DB 1; Length 602;
DB	Best Local Similarity		73.1%;	Pred. No. 1.6e-121;
DB	Matches	444;	Conservative	55; Mismatches 66; Indels 42; Gaps 5;
QY	69	ESAGCSQSPALPPRLKEMKKSQESAGSKVLRCETSSSEYSLRPFKFWKNGHNLKKNKPQ	128	
DB	17	QTDVNSVNTVPKLNKRNORVAVGQKLVLCRTTSEYPALRFKLNKNGEITKKNRPE	76	
QY	129	NIKTQKPGK-SELIRINKASLADSGEYMKYISKLGNDSTKASVITDTN-----	187	
DB	77	NVKIPKOKKYSELHIYRATLADAGEACRVSSKLGNDSTKASVITDTN-----	126	
QY	188	EGATVSESPTRISVSTEGANTSSSTSTGTGTHLVKCAKEKTCFVNGGECFMVKDLS	247	
DB	127	-----ATSTSTGTGTHLVKCAKEKTCFVNGGECFMVKDLS	162	
QY	248	NPSRYLCKCPGFTCARCTENVPKMQNQ-----EKAELYOKRVLTITGICIALLVVGI	302	
DB	163	NPRLYLCPCNEFTGDRCONVYMASFYKHLGIEFMEAEELYOKRVLTITGICIALLVVGI	222	
QY	303	MCVAYCKTKQRKLLHDLRQLRSERNNMANGPHHPNPPNVQVNVQVSKNVI	362	
DB	223	MCVAYCKTKQRKLLHDLRQLRSERNNMANGPHHPNPPNVQVNVQVSKNII	282	
QY	363	SSEHIVERAETSSTSHYTSTAHTSTVTTPSHSWSNGHTEILSESHSVIVSVSN	422	
DB	283	SSERVVERETSTSTSHYTSTTHHSMVTVTQPSHSWSNGHTEILSESHSVIVSVSN	342	
QY	423	SRHSSTPGPRGLNGTGGPRECNFLRHARETDPDSYEDSPHSERYVSAMTTPARMSPD	482	
DB	343	SRHTSPT-GPRGLNGIGGPREGNSFLRHARETDPDSYEDSPHSERYVSAMTTPARMSPD	401	
QY	483	FHTSSPKSPSEMSPPVSSMTVSNMSPMAVSPFMEERPLLVTPPRLREKFFDHPQOF	542	
DB	402	FHTSPSPKSPSEMSPPVSSMTVSNMSPMAVSPFMEERPLLVTPPRLRE-KYDNHQQOF	460	
QY	543	SSFHNPADNSLPASPLRIVEDEYETTOYEYEPQAQEPVKKLANSRKRTKPNGHIAN	602	
DB	461	NSFHNNPHTSNLPPSPPLRIVEDEYETTOYEYEPQAQEPVKKLANSRKRTKPNGHIAN	520	
QY	603	RLEVDNYSOSSNSESETERVEDPTFFGIQNPILAASLEATPAFLADSRTPNAGRF	662	

Db 351 TSFSTSHVTSHTSHSTVTTQTPSHSWSNGLSBSMISEKSYSVIVTSSVENSHTSPT-GP 409
 QY 433 RGRNGTGPRCNSFLRHARTPDSDYRSPHSERYVSAMTTPARMSPVDFHTPPSPKSP 492
 Db 410 RGRNGIGPRDC-SYLHARDTPDSYRSPHSERYVSAMTTPARMSPVEFKTPIPSKSP 468
 QY 493 PSEMPYPVSTVMSNAPSVMFMEERPLLVTPRLREKFDH-PQ-----QFSSEH 546
 Db 469 CUETSPSSSLAVSVFSAVSFIEERPLLVSPRLREKRYDRKTPKPHKQNSYH 528
 QY 547 HNPADHSNLSPLRIVEDEYETQYPAQEPVKKLANSRRAKRTKPNCHIANRLEV 606
 Db 529 HNPCHDSSSLPNLRIVEDEYETQYEPVLEPAKLVNSRRKRTKPNCHISNREL 588
 QY 607 DNTSSQSSNSSETEDEKVGDTDFLQNPPLAASLEATPAFLADSRNTPAGRFSTQE 666
 Db 589 DSDSSSESTSETEDEKVGDTDFLQNPPLAASLEASLYRADSRNTPSRTSTQE 648
 QY 667 EQ 669
 649 EQ 651

RESULT 5

NRG2_HUMAN
 ID NRG2_HUMAN STANDARD; PRT; 850 AA.
 AC O14511;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pro-neuregulin-2 precursor (Pro-NRG2) [Contains: Neuregulin-2 (NRG-2)
 DE (Neural- and thymus-derived activator for ERBB kinases) (NTAK)
 DE (Divergent of neuregulin 1) (DON-1)].
 GN NRG2 OR NTAK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Neuroblastoma;
 RX MEDLINE=98006324; PubMed=9348101;
 RA Higashiyama S., Horikawa M., Yamada K., Ichino N., Nakano N.,
 RA Nakagawa T., Miyagawa J., Matsushita N., Nagatsu T., Taniguchi N.,
 RA Ishiguro H.;
 RT "A novel brain-derived member of the epidermal growth factor family
 RT that interacts with ErbB3 and ErbB4.";
 J. Biochem. 122:675-680(1997).
 [2]
 RP SEQUENCE FROM N.A. (ISOFORMS DON-1B AND DON-1R).
 RC TISSUE=Fetal brain;
 RX MEDLINE=97342638; PubMed=9199335;
 RA Busfield S.J., Michnick D.A., Chickering T.W., Revett T.L., Ma J.,
 RA Woolf E.A., Comrack C.A., Dussault B.J., Woolf J., Goodearl A.D.J.,
 RA Gearing D.P.;
 RT "Characterization of a neuregulin-related gene, Don-1, that is highly
 RT expressed in restricted regions of the cerebellum and hippocampus.";
 RL Mol. Cell. Biol. 17:4007-4014(1997).
 [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6).
 RC TISSUE=Lung, and Fetal brain;
 RX MEDLINE=99295836; PubMed=10369162;
 RA Ring H.Z., Chang H., Guilbot A., Brice A., LeGuern E., Francke U.;
 RT "The human neuregulin 2 (NRG2) gene: cloning, mapping and evaluation
 RT as a candidate for the autosomal recessive form of Charcot-Marie-Tooth
 RT disease linked to 5q.";
 RL Hum. Genet. 104:326-332(1999).
 [3]
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE
 CC HETERODIMERIZATION WITH THE EGF RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS

A PROTEOLYTICALLY RELEASED SOLUBLE GROWTH FACTOR FORM. THE
 MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: 8 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4, 5, 6,
 CC DON-1B AND DON-1R; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO THE CEREBELLUM IN THE ADULT.
 CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION
 CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE
 CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN
 CC DIMERIZATION (BY SIMILARITY).
 CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE
 CC DOMAIN (BY SIMILARITY).
 CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE
 CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR
 CC FORM (BY SIMILARITY).
 CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB005060; BAA23417.1;
 CC EMBL; AF119162; AAF28848.1;
 CC EMBL; AF119151; AAF28848.1; JOINED.
 CC EMBL; AF119151; AAF28848.1; JOINED.
 CC EMBL; AF119152; AAF28848.1; JOINED.
 CC EMBL; AF119153; AAF28848.1; JOINED.
 CC EMBL; AF119154; AAF28848.1; JOINED.
 CC EMBL; AF119155; AAF28848.1; JOINED.
 CC EMBL; AF119156; AAF28848.1; JOINED.
 CC EMBL; AF119157; AAF28848.1; JOINED.
 CC EMBL; AF119158; AAF28848.1; JOINED.
 CC EMBL; AF119159; AAF28848.1; JOINED.
 CC EMBL; AF119160; AAF28848.1; JOINED.
 CC EMBL; AF119161; AAF28848.1; JOINED.
 CC EMBL; AF119162; AAF28848.1; JOINED.
 CC EMBL; AF119163; AAF28848.1; JOINED.
 CC EMBL; AF119164; AAF28848.1; JOINED.
 CC EMBL; AF119165; AAF28848.1; JOINED.
 CC EMBL; AF119166; AAF28848.1; JOINED.
 CC EMBL; AF119167; AAF28848.1; JOINED.
 CC EMBL; AF119168; AAF28848.1; JOINED.
 CC EMBL; AF119169; AAF28848.1; JOINED.
 CC EMBL; AF119170; AAF28848.1; JOINED.
 CC EMBL; AF119171; AAF28848.1; JOINED.
 CC EMBL; AF119172; AAF28848.1; JOINED.
 CC EMBL; AF119173; AAF28848.1; JOINED.
 CC EMBL; AF119174; AAF28848.1; JOINED.
 CC EMBL; AF119175; AAF28848.1; JOINED.
 CC EMBL; AF119176; AAF28848.1; JOINED.
 CC EMBL; AF119177; AAF28848.1; JOINED.
 CC EMBL; AF119178; AAF28848.1; JOINED.
 CC EMBL; AF119179; AAF28848.1; JOINED.
 CC EMBL; AF119180; AAF28848.1; JOINED.
 CC EMBL; AF119181; AAF28848.1; JOINED.
 CC EMBL; AF119182; AAF28848.1; JOINED.
 CC EMBL; AF119183; AAF28848.1; JOINED.
 CC EMBL; AF119184; AAF28848.1; JOINED.
 CC EMBL; AF119185; AAF28848.1; JOINED.
 CC EMBL; AF119186; AAF28848.1; JOINED.
 CC EMBL; AF119187; AAF28848.1; JOINED.
 CC EMBL; AF119188; AAF28848.1; JOINED.
 CC EMBL; AF119189; AAF28848.1; JOINED.
 CC EMBL; AF119190; AAF28848.1; JOINED.
 CC EMBL; AF119191; AAF28848.1; JOINED.
 CC EMBL; AF119192; AAF28848.1; JOINED.
 CC EMBL; AF119193; AAF28848.1; JOINED.
 CC EMBL; AF119194; AAF28848.1; JOINED.
 CC EMBL; AF119195; AAF28848.1; JOINED.
 CC EMBL; AF119196; AAF28848.1; JOINED.
 CC EMBL; AF119197; AAF28848.1; JOINED.
 CC EMBL; AF119198; AAF28848.1; JOINED.
 CC EMBL; AF119199; AAF28848.1; JOINED.
 CC EMBL; AF119200; AAF28848.1; JOINED.
 CC EMBL; AF119201; AAF28848.1; JOINED.
 CC EMBL; AF119202; AAF28848.1; JOINED.
 CC EMBL; AF119203; AAF28848.1; JOINED.
 CC EMBL; AF119204; AAF28848.1; JOINED.
 CC EMBL; AF119205; AAF28848.1; JOINED.
 CC EMBL; AF119206; AAF28848.1; JOINED.
 CC EMBL; AF119207; AAF28848.1; JOINED.
 CC EMBL; AF119208; AAF28848.1; JOINED.
 CC EMBL; AF119209; AAF28848.1; JOINED.
 CC EMBL; AF119210; AAF28848.1; JOINED.
 CC EMBL; AF119211; AAF28848.1; JOINED.
 CC EMBL; AF119212; AAF28848.1; JOINED.
 CC EMBL; AF119213; AAF28848.1; JOINED.
 CC EMBL; AF119214; AAF28848.1; JOINED.
 CC EMBL; AF119215; AAF28848.1; JOINED.
 CC EMBL; AF119216; AAF28848.1; JOINED.
 CC EMBL; AF119217; AAF28848.1; JOINED.
 CC EMBL; AF119218; AAF28848.1; JOINED.
 CC EMBL; AF119219; AAF28848.1; JOINED.
 CC EMBL; AF119220; AAF28848.1; JOINED.
 CC EMBL; AF119221; AAF28848.1; JOINED.
 CC EMBL; AF119222; AAF28848.1; JOINED.
 CC EMBL; AF119223; AAF28848.1; JOINED.
 CC EMBL; AF119224; AAF28848.1; JOINED.
 CC EMBL; AF119225; AAF28848.1; JOINED.
 CC EMBL; AF119226; AAF28848.1; JOINED.
 CC EMBL; AF119227; AAF28848.1; JOINED.
 CC EMBL; AF119228; AAF28848.1; JOINED.
 CC EMBL; AF119229; AAF28848.1; JOINED.
 CC EMBL; AF119230; AAF28848.1; JOINED.
 CC EMBL; AF119231; AAF28848.1; JOINED.
 CC EMBL; AF119232; AAF28848.1; JOINED.
 CC EMBL; AF119233; AAF28848.1; JOINED.
 CC EMBL; AF119234; AAF28848.1; JOINED.
 CC EMBL; AF119235; AAF28848.1; JOINED.
 CC EMBL; AF119236; AAF28848.1; JOINED.
 CC EMBL; AF119237; AAF28848.1; JOINED.
 CC EMBL; AF119238; AAF28848.1; JOINED.
 CC EMBL; AF119239; AAF28848.1; JOINED.
 CC EMBL; AF119240; AAF28848.1; JOINED.
 CC EMBL; AF119241; AAF28848.1; JOINED.
 CC EMBL; AF119242; AAF28848.1; JOINED.
 CC EMBL; AF119243; AAF28848.1; JOINED.
 CC EMBL; AF119244; AAF28848.1; JOINED.
 CC EMBL; AF119245; AAF28848.1; JOINED.
 CC EMBL; AF119246; AAF28848.1; JOINED.
 CC EMBL; AF119247; AAF28848.1; JOINED.
 CC EMBL; AF119248; AAF28848.1; JOINED.
 CC EMBL; AF119249; AAF28848.1; JOINED.
 CC EMBL; AF119250; AAF28848.1; JOINED.
 CC EMBL; AF119251; AAF28848.1; JOINED.
 CC EMBL; AF119252; AAF28848.1; JOINED.
 CC EMBL; AF119253; AAF28848.1; JOINED.
 CC EMBL; AF119254; AAF28848.1; JOINED.
 CC EMBL; AF119255; AAF28848.1; JOINED.
 CC EMBL; AF119256; AAF28848.1; JOINED.
 CC EMBL; AF119257; AAF28848.1; JOINED.
 CC EMBL; AF119258; AAF28848.1; JOINED.
 CC EMBL; AF119259; AAF28848.1; JOINED.
 CC EMBL; AF119260; AAF28848.1; JOINED.
 CC EMBL; AF119261; AAF28848.1; JOINED.
 CC EMBL; AF119262; AAF28848.1; JOINED.
 CC EMBL; AF119263; AAF28848.1; JOINED.
 CC EMBL; AF119264; AAF28848.1; JOINED.
 CC EMBL; AF119265; AAF28848.1; JOINED.
 CC EMBL; AF119266; AAF28848.1; JOINED.
 CC EMBL; AF119267; AAF28848.1; JOINED.
 CC EMBL; AF119268; AAF28848.1; JOINED.
 CC EMBL; AF119269; AAF28848.1; JOINED.
 CC EMBL; AF119270; AAF28848.1; JOINED.
 CC EMBL; AF119271; AAF28848.1; JOINED.
 CC EMBL; AF119272; AAF28848.1; JOINED.
 CC EMBL; AF119273; AAF28848.1; JOINED.
 CC EMBL; AF119274; AAF28848.1; JOINED.
 CC EMBL; AF119275; AAF28848.1; JOINED.
 CC EMBL; AF119276; AAF28848.1; JOINED.
 CC EMBL; AF119277; AAF28848.1; JOINED.
 CC EMBL; AF119278; AAF28848.1; JOINED.
 CC EMBL; AF119279; AAF28848.1; JOINED.
 CC EMBL; AF119280; AAF28848.1; JOINED.
 CC EMBL; AF119281; AAF28848.1; JOINED.
 CC EMBL; AF119282; AAF28848.1; JOINED.
 CC EMBL; AF119283; AAF28848.1; JOINED.
 CC EMBL; AF119284; AAF28848.1; JOINED.
 CC EMBL; AF119285; AAF28848.1; JOINED.
 CC EMBL; AF119286; AAF28848.1; JOINED.
 CC EMBL; AF119287; AAF28848.1; JOINED.
 CC EMBL; AF119288; AAF28848.1; JOINED.
 CC EMBL; AF119289; AAF28848.1; JOINED.
 CC EMBL; AF119290; AAF28848.1; JOINED.
 CC EMBL; AF119291; AAF28848.1; JOINED.
 CC EMBL; AF119292; AAF28848.1; JOINED.
 CC EMBL; AF119293; AAF28848.1; JOINED.
 CC EMBL; AF119294; AAF28848.1; JOINED.
 CC EMBL; AF119295; AAF28848.1; JOINED.
 CC EMBL; AF119296; AAF28848.1; JOINED.
 CC EMBL; AF119297; AAF28848.1; JOINED.
 CC EMBL; AF119298; AAF28848.1; JOINED.
 CC EMBL; AF119299; AAF28848.1; JOINED.
 CC EMBL; AF119300; AAF28848.1; JOINED.
 CC EMBL; AF119301; AAF28848.1; JOINED.
 CC EMBL; AF119302; AAF28848.1; JOINED.
 CC EMBL; AF119303; AAF28848.1; JOINED.
 CC EMBL; AF119304; AAF28848.1; JOINED.
 CC EMBL; AF119305; AAF28848.1; JOINED.
 CC EMBL; AF119306; AAF28848.1; JOINED.
 CC EMBL; AF119307; AAF28848.1; JOINED.
 CC EMBL; AF119308; AAF28848.1; JOINED.
 CC EMBL; AF119309; AAF28848.1; JOINED.
 CC EMBL; AF119310; AAF28848.1; JOINED.
 CC EMBL; AF119311; AAF28848.1; JOINED.
 CC EMBL; AF119312; AAF28848.1; JOINED.
 CC EMBL; AF119313; AAF28848.1; JOINED.
 CC EMBL; AF119314; AAF28848.1; JOINED.
 CC EMBL; AF119315; AAF28848.1; JOINED.
 CC EMBL; AF119316; AAF28848.1; JOINED.
 CC EMBL; AF119317; AAF28848.1; JOINED.
 CC EMBL; AF119318; AAF28848.1; JOINED.
 CC EMBL; AF119319; AAF28848.1; JOINED.
 CC EMBL; AF119320; AAF28848.1; JOINED.
 CC EMBL; AF119321; AAF28848.1; JOINED.
 CC EMBL; AF119322; AAF28848.1; JOINED.
 CC EMBL; AF119323; AAF28848.1; JOINED.
 CC EMBL; AF119324; AAF28848.1; JOINED.
 CC EMBL; AF119325; AAF28848.1; JOINED.
 CC EMBL; AF119326; AAF28848.1; JOINED.
 CC EMBL; AF119327; AAF28848.1; JOINED.
 CC EMBL; AF119328; AAF28848.1; JOINED.
 CC EMBL; AF119329; AAF28848.1; JOINED.
 CC EMBL; AF119330; AAF28848.1; JOINED.
 CC EMBL; AF119331; AAF28848.1; JOINED.
 CC EMBL; AF119332; AAF28848.1; JOINED.
 CC EMBL; AF119333; AAF28848.1; JOINED.
 CC EMBL; AF119334; AAF28848.1; JOINED.
 CC EMBL; AF119335; AAF28848.1; JOINED.
 CC EMBL; AF119336; AAF28848.1; JOINED.
 CC EMBL; AF119337; AAF28848.1; JOINED.
 CC EMBL; AF119338; AAF28848.1; JOINED.
 CC EMBL; AF119339; AAF28848.1; JOINED.
 CC EMBL; AF119340; AAF28848.1; JOINED.
 CC EMBL; AF119341; AAF28848.1; JOINED.
 CC EMBL; AF119342; AAF28848.1; JOINED.
 CC EMBL; AF119343; AAF28848.1; JOINED.
 CC EMBL; AF119344; AAF28848.1; JOINED.
 CC EMBL; AF119345; AAF28848.1; JOINED.
 CC EMBL; AF119346; AAF28848.1; JOINED.
 CC EMBL; AF119347; AAF28848.1; JOINED.
 CC EMBL; AF119348; AAF28848.1; JOINED.
 CC EMBL; AF119349; AAF28848.1; JOINED.
 CC EMBL; AF119350; AAF28848.1; JOINED.
 CC EMBL; AF119351; AAF28848.1; JOINED.
 CC EMBL; AF119352; AAF28848.1; JOINED.
 CC EMBL; AF119353; AAF28848.1; JOINED.
 CC EMBL; AF119354; AAF28848.1; JOINED.
 CC EMBL; AF119355; AAF28848.1; JOINED.
 CC EMBL; AF119356; AAF28848.1; JOINED.
 CC EMBL; AF119357; AAF28848.1; JOINED.
 CC EMBL; AF119358; AAF28848.1; JOINED.
 CC EMBL; AF119359; AAF28848.1; JOINED.
 CC EMBL; AF119360; AAF28848.1; JOINED.
 CC EMBL; AF119361; AAF28848.1; JOINED.
 CC EMBL; AF119362; AAF28848.1; JOINED.
 CC EMBL; AF119363; AAF28848.1; JOINED.
 CC EMBL; AF119364; AAF28848.1; JOINED.
 CC EMBL; AF119365; AAF28848.1; JOINED.
 CC EMBL; AF119366; AAF28848.1; JOINED.
 CC EMBL; AF119367; AAF28848.1; JOINED.
 CC EMBL; AF119368; AAF28848.1; JOINED.
 CC EMBL; AF119369; AAF28848.1; JOINED.
 CC EMBL; AF119370; AAF28848.1; JOINED.
 CC EMBL; AF119371; AAF28848.1; JOINED.
 CC EMBL; AF119372; AAF28848.1; JOINED.
 CC EMBL; AF119373; AAF28848.1; JOINED.
 CC EMBL; AF119374; AAF28848.1; JOINED.
 CC EMBL; AF119375; AAF28848.1; JOINED.
 CC EMBL; AF119376; AAF28848.1; JOINED.
 CC EMBL; AF119377; AAF28848.1; JOINED.
 CC EMBL; AF119378; AAF28848.1; JOINED.
 CC EMBL; AF119379; AAF28848.1; JOINED.
 CC EMBL; AF119380; AAF28848.1; JOINED.
 CC EMBL; AF119381; AAF28848.1; JOINED.
 CC EMBL; AF119382; AAF28848.1; JOINED.
 CC EMBL; AF119383; AAF28848.1; JOINED.
 CC EMBL; AF119384; AAF28848.1; JOINED.
 CC EMBL; AF119385; AAF28848.1; JOINED.
 CC EMBL; AF119386; AAF28848.1; JOINED.
 CC EMBL; AF119387; AAF28848.1; JOINED.
 CC EMBL; AF119388; AAF28848.1; JOINED.
 CC EMBL; AF119389; AAF28848.1; JOINED.
 CC EMBL; AF119390; AAF28848.1; JOINED.
 CC EMBL; AF119391; AAF28848.1; JOINED.
 CC EMBL; AF119392; AAF28848.1; JOINED.
 CC EMBL; AF119393; AAF28848.1; JOINED.
 CC EMBL; AF119394; AAF28848.1; JOINED.
 CC EMBL; AF119395; AAF28848.1; JOINED.
 CC EMBL; AF119396; AAF28848.1; JOINED.
 CC EMBL; AF119397; AAF28848.1; JOINED.
 CC EMBL; AF119398; AAF28848.1; JOINED.
 CC EMBL; AF119399; AAF28848.1; JOINED.
 CC EMBL; AF119400; AAF28848.1; JOINED.
 CC EMBL; AF119401; AAF28848.1; JOINED.
 CC EMBL; AF119402; AAF28848.1; JOINED.
 CC EMBL; AF119403; AAF28848.1; JOINED.
 CC EMBL; AF119404; AAF28848.1; JOINED.
 CC EMBL; AF119405; AAF28848.1; JOINED.
 CC EMBL; AF119406; AAF28848.1; JOINED.
 CC EMBL; AF119407; AAF28848.1; JOINED.
 CC EMBL; AF119408; AAF28848.1; JOINED.
 CC EMBL; AF119409; AAF28848.1; JOINED.
 CC EMBL; AF119410; AAF28848.1; JOINED.
 CC EMBL; AF119411; AAF28848.1; JOINED.
 CC EMBL; AF119412; AAF28848.1; JOINED.
 CC EMBL; AF119413; AAF28848.1; JOINED.
 CC EMBL; AF119414; AAF28848.1; JOINED.
 CC EMBL; AF119415; AAF28848.1; JOINED.
 CC EMBL; AF119416; AAF28848.1; JOINED.
 CC EMBL; AF119417; AAF28848.1; JOINED.
 CC EMBL; AF119418; AAF28848.1; JOINED.
 CC EMBL; AF119419; AAF28848.1; JOINED.
 CC EMBL; AF119420; AAF28848.1; JOINED.
 CC EMBL; AF119421; AAF28848.1; JOINED.
 CC EMBL; AF119422; AAF28848.1; JOINED.
 CC EMBL; AF119423; AAF28848.1; JOINED.
 CC EMBL; AF119424; AAF28848.1; JOINED.
 CC EMBL; AF119425; AAF28848.1; JOINED.
 CC EMBL; AF119426; AAF28848.1; JOINED.
 CC EMBL; AF119427; AAF28848.1; JOINED.
 CC EMBL; AF119428; AAF28848.1; JOINED.
 CC EMBL; AF119429; AAF28848.1; JOINED.
 CC EMBL; AF119430; AAF28848.1; JOINED.
 CC EMBL; AF119431; AAF28848.1; JOINED.
 CC EMBL; AF119432; AAF28848.1; JOINED.
 CC EMBL; AF119433; AAF28848.1; JOINED.
 CC EMBL; AF119434; AAF28848.1; JOINED.
 CC EMBL; AF119435; AAF28848.1; JOINED.
 CC EMBL; AF119436; AAF28848.1; JOINED.
 CC EMBL; AF119437; AAF28848.1; JOINED.
 CC EMBL; AF119438; AAF28848.1; JOINED.
 CC EMBL; AF119439; AAF28848.1; JOINED.
 CC EMBL; AF119440; AAF28848.1; JOINED.
 CC EMBL; AF119441; AAF28848.1; JOINED.
 CC EMBL; AF119442; AAF28848.1; JOINED.
 CC EMBL; AF119443; AAF28848.1; JOINED.
 CC EMBL; AF119444; AAF28848.1; JOINED.
 CC EMBL; AF119445; AAF28848.1; JOINED.
 CC EMBL; AF119446; AAF28848.1; JOINED.
 CC EMBL; AF119447; AAF28848.1; JOINED.
 CC EMBL; AF119448; AAF28848.1; JOINED.
 CC EMBL; AF119449; AAF28848.1; JOINED.
 CC EMBL; AF119450; AAF28848.1; JOINED.
 CC EMBL; AF119451; AAF28848.1; JOINED.
 CC EMBL; AF119452; AAF28848.1; JOINED.
 CC EMBL; AF119453; AAF28848.1; JOINED.
 CC EMBL; AF119454; AAF28848.1; JOINED.
 CC EMBL; AF119455; AAF28848.1; JOINED.
 CC EMBL; AF119456; AAF28848.1; JOINED.
 CC EMBL; AF119457; AAF28848.1; JOINED.
 CC EMBL; AF119458; AAF28848.1; JOINED.
 CC EMBL; AF119459; AAF28848.1; JOINED.
 CC EMBL; AF119460; AAF28848.1; JOINED.
 CC EMBL; AF119461; AAF28848.1; JOINED.
 CC EMBL; AF119462; AAF28848.1; JOINED.
 CC EMBL; AF119463; AAF28848.1; JOINED.
 CC EMBL; AF119464; AAF28848.1; JOINED.
 CC EMBL; AF119465; AAF28848.1; JOINED.
 CC EMBL; AF119466; AAF28848.1; JOINED.
 CC EMBL; AF119467; AAF28848.1; JOINED.
 CC EMBL; AF119468; AAF28848.1; JOINED.
 CC EMBL; AF119469; AAF28848.1; JOINED.
 CC EMBL; AF119470; AAF28848.1; JOINED.
 CC EMBL; AF119471; AAF28848.1; JOINED.
 CC EMBL; AF119472; AAF28848.1; JOINED.
 CC EMBL; AF119473; AAF28848.1; JOINED.
 CC EMBL; AF119474; AAF28848.1; JOINED.
 CC EMBL; AF119475; AAF28848.1; JOINED.
 CC EMBL; AF119476; AAF28848.1; JOINED.
 CC EMBL; AF119477; AAF28848.1; JOINED.
 CC EMBL; AF119478; AAF28848.1; JOINED.
 CC EMBL; AF119479; AAF28848.1; JOINED.
 CC EMBL; AF119480; AAF28848.1; JOINED.

RA Ishiguro H.:

RT "A novel brain-derived member of the epidermal growth factor family

RL that interacts with ErbB3 and ErbB4.";

RN J. Biochem. 122:675-680(1997).

CC [2]

CC SEQUENCE OF 109-868 FROM N.A. (ISOFORMS NR2-ALPHA AND NR2-BETA).

CC TISSUE-Cerebellum;

CC MEDLINE-97311397; PubMed-9168114;

CC Chang H., Riese D.J. II, Gilbert W., Stern D.F., McMahan U.J.;

CC "Ligands for ErbB-family receptors encoded by a neuregulin-like

CC gene.";

CC Nature 387:509-512(1997).

CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE

CC RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORRECEPTORS,

CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND

CC ACTIVATION OF THE ERBB RECEPTORS. MAY ALSO PROMOTE THE

CC HETERODIMERIZATION WITH THE EGF RECEPTOR.

CC -1- SUBCELLULAR LOCATION: EXISTS AS AN TYPE I MEMBRANE PROTEIN AND AS

CC MEMBRANE-BOUND FORM DOES NOT SEEM TO BE ACTIVE (BY SIMILARITY).

CC ALTERNATIVE PRODUCTS: AT LEAST 7 ISOFORMS; NTA-ALPHA1 (SHOWN

CC HERE), NTA-ALPHA2A, NTA-ALPHA2B/NTAK-ALPHA2-1P, NTA-ALPHA2, NTA-ALPHA2-1P,

CC GAMMA, NR2-ALPHA AND NR2-BETA; ARE PRODUCED BY ALTERNATIVE

CC SPLICING. THE ALPHA-TYPE AND BETA-TYPE DIFFER IN THE EGF-LIKE

CC DOMAIN.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN MOST PARTS OF THE BRAIN,

CC ESPECIALLY THE OLFACTORY BULB AND CEREBELLUM WHERE IT LOCALIZES IN

CC GRANULE AND PURKINJE CELLS. IN THE HIPPOCAMPUS, FOUND IN THE

CC GRANULE CELLS OF THE DENTATE GYRUS. IN THE BASAL FOREBRAIN, FOUND

CC IN THE CHOLINERGIC CELLS. IN THE HINDRAIN, WEAKLY DETECTABLE IN

CC THE MOTOR TRIGEMINAL NUCLEUS. NOT DETECTED IN THE HYPOTHALAMUS.

CC ALSO FOUND IN THE LIVER AND IN THE THYMUS. NOT DETECTED IN HEART,

CC ADRENAL GLAND, OR TESTIS.

CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, EXPRESSED IN THE BRAIN OF

CC E11.5 EMBRYOS WHERE IT IS FOUND IN THE TELENCEPHALON, BUT NOT IN

CC THE HINDRAIN. NOT FOUND IN THE HEART. IN THE ADULT, FOUND IN

CC BRAIN AND THYMUS

CC -1- DOMAIN: THE CYTOPLASMIC DOMAIN MAY BE INVOLVED IN THE REGULATION

CC OF TRAFFICKING AND PROTEOLYTIC PROCESSING. REGULATION OF THE

CC PROTEOLYTIC PROCESSING INVOLVES INITIAL INTRACELLULAR DOMAIN

CC DIMERIZATION (BY SIMILARITY).

CC -1- DOMAIN: ERBB RECEPTOR BINDING IS ELICITED ENTIRELY BY THE EGF-LIKE

CC DOMAIN (BY SIMILARITY).

CC -1- PTM: PROTEOLYTIC CLEAVAGE CLOSE TO THE PLASMA MEMBRANE ON THE

CC EXTERNAL FACE LEADS TO THE RELEASE OF THE SOLUBLE GROWTH FACTOR

CC FORM (BY SIMILARITY).

CC -1- PTM: EXTENSIVE GLYCOSYLATION PRECEDES THE PROTEOLYTIC CLEAVAGE (BY

CC SIMILARITY).

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: D89995; BAA23344.1; -

CC EMBL: D89996; BAA23345.1; -

CC EMBL: D89997; BAA23346.1; -

CC EMBL: D89998; BAA23347.1; -

CC EMBL: AB001576; BAA23348.1; -

CC HSSP: Q12784; 1HRE.

CC InterPro: IPR000561; EGF-like.

CC InterPro: IPR003006; Ig_MHC.

CC InterPro: IPR003598; Ig_c2.

CC InterPro: IPR002154; Neuregulin.

CC Pfam: PF00008; EGF; 1.

CC Pfam: PF00047; Ig; 1.

CC Pfam: PF02158; Neuregulin; 1.

DR SMART; SM00181; EGF; 1.

DR SMART; SM00408; IGC2; 1.

DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; 1.

KW Growth factor; EGF-like domain; Immunoglobulin domain; Glycoprotein;

KW Transmembrane; Multigene family; Alternative splicing.

FT PROPEP 1 127

FT CHAIN 128 868 PRO-NEUREGULIN-2, MEMBRANE-BOUND FORM.

FT CHAIN 128 868 NEUREGULIN-2.

FT DOMAIN 128 428 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 430 450 INTERNAL SIGNAL SEQUENCE (POTENTIAL).

FT DOMAIN 451 868 CYTOPLASMIC (POTENTIAL).

FT DOMAIN 266 334 IG-LIKE C2-TYPE DOMAIN.

FT DOMAIN 336 398 SER/THR-RICH.

FT DOMAIN 357 398 EGF-LIKE.

FT DOMAIN 22 32 POLY-SER.

FT DOMAIN 35 45 POLY-THR.

FT DOMAIN 56 59 POLY-ALA.

FT DOMAIN 103 106 POLY-PRO.

FT DISULFID 273 327 BY SIMILARITY.

FT DISULFID 361 375 BY SIMILARITY.

FT DISULFID 369 386 BY SIMILARITY.

FT DISULFID 388 397 BY SIMILARITY.

FT CARBOHYD 33 33 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 1 108 MISSING (IN ISOFORM NR2-BETA).

FT VARSPLIC 220 222 PLV -> FFF (IN ISOFORM NTA-ALPHA2-1P).

FT VARSPLIC 388 388 MISSING (IN ISOFORM NTA-ALPHA2-1P).

FT VARSPLIC 390 412 NGFFQRCLEKLPRLYMPDPKQ -> VGYTDRCOQFAMV

FT VARSPLIC 390 421 NFFS (IN ISOFORM NR2-BETA).

FT VARSPLIC 414 421 NGFFQRCLEKLPRLYMPDPKQKHLGFELKE -> VGYTDRCOQFAMV

FT VARSPLIC 414 439 MISSING (IN ISOFORM NTA-ALPHA2A AND ISOFORM NTA-ALPHA2B).

FT VARSPLIC 440 868 HLGFEKLEAEYQKRVLTIGTICVA -> SVLWDPGPGV

FT CONFLICT 117 117 S -> F (IN REF. 2).

FT CONFLICT 724 724 R -> H (IN REF. 2).

SQ SEQUENCE 868 AA; 93776 MW; 3C7D4D94DBE64DE2 CRC64;

Query Match 28.9%; Score 1001.5; DB 1; Length 868;

Best Local Similarity 36.8%; Pred. No. 1.6e-48;

Matches 248; Conservative 102; Mismatches 184; Indels 139; Gaps 21;

QY 55 GKGKKKRGSGKKPSAAGSQSPALPPRLKMKMSQESAAGSKVLRCETSSYSLRFRW 114

DB 234 GKNIRKEVGKILTCAT-----RPKLKMKSGTGEVGEKQSLKCEAAAGNPQPSYRW 286

QY 115 FNGNELAKNKPNQIKKPKGK--SELINKASLADSGEYMKVSKLGNDSASANT 172

DB 287 FDKGKELNRS-----RDRIKYGKRNKSRLOFNKRVKEDAGEYVCEAENLIGDVTYGRUH 343

QY 173 IVESNEITGMPASTGAYVSSPISVSTEGANTSSSTSTSTGTSHLVKCAEKEKT 232

DB 344 V-----NSVSTTSSWSG--HARKCNETAKS 367

QY 233 FCVNGGECFMVKDLNPNRYLCKCPGFTGARTENVPKVO-----NOEKABE 281

DB 368 YCVNGVCYVIEGINQLS---CKCPNGPFGQCLEKLPRLYMPDPKQKHLGFELKEABE 424

QY 282 LYQKRVLTITGICIAALLVVGIMCVAYCKTKQKRLHDLRQLRSERNRMNANGPH 341

DB 425 LYQKRVLTITGICVALLVVGIMCVAYCKTKQKRLHDLRQLRSERNRMNANGPH 483

QY 342 HNPPEPVOLVNOYVKNVISSEHIVERAEATSFSTSHYTTAHSHTTDTOT-----PS 396

DB 484 HPRLDEEIQMAD--YISKKNVPATDHYIRREAEATTFSGSHSCSPSHHCSTATPTSSHRHES 542

ID AC Q15491, STANDARD; PRT; 296 AA.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Neuregulin-1, sensory and motor neuron-derived factor isoform.
 GN NRGL OR HCL OR NDF OR HRGA OR GGF OR SMDF.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 ON NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain stem, and Cerebellum;
 RA MEDLINE=95301541; PubMed=7782315;
 RT HO W.-H., Armanini M.P., Nijfens A., Phillips H.S., Osheroff P.L.;
 RT "Sensory and motor neuron-derived factor. A novel heregulin variant
 RL highly expressed in sensory and motor neurons.";
 RL J. Biol. Chem. 270:14523-14532(1995).
 CC -1- FUNCTION: THE ISOFORM SMDF MAY PLAY A ROLE IN MOTOR AND SENSORY
 CC NEURON DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: SECRETED. MAY POSSESS AN INTERNAL UNCLEAVED
 CC SIGNAL SEQUENCE.
 CC -1- ALTERNATIVE PRODUCTS: At least 10 isoforms: alpha (AC Q02297);
 CC alpha1A, alpha2B, alpha3, beta1, beta2, beta3/GGFHEB1,
 CC GGF2/GGFHP2 and SMDF (shown here); are produced by alternative
 CC splicing. They have been classified as type I NRGs (variants with
 CC an IG domain and a glycosylation domain; alpha and beta), Type II
 CC NRGs (variants with an IG domain but no glycosylation domain;
 CC GGF2) and type III NRGs (variants with a cysteine-rich domain; SMDF).
 CC All these isoforms perform distinct tissue-specific functions.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM: SPINAL CORD MOTOR
 CC NEURONS, DORSAL ROOT GANGLION NEURONS, AND BRAIN. PREDOMINANT
 CC ISOFORM EXPRESSED IN SENSORY AND MOTOR NEURONS. NOT DETECTED IN
 CC ADULT HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND
 CC PANCREAS. NOT EXPRESSED IN FETAL LUNG, LIVER, AND KIDNEY.
 CC -1- DEVELOPMENTAL STAGE: HIGHLY EXPRESSED IN DEVELOPING SPINAL MOTOR
 CC NEURONS AND IN DEVELOPING CRANIAL NERVE NUCLEI. EXPRESSION IS
 CC MAINTAINED ONLY IN BOTH ADULT MOTOR NEURONS AND DORSAL ROOT
 CC GANGLION NEURONS.
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NEUREGULIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L41827; AAC1764.1; --
 CC HSSP; Q12784; IHRF.
 CC Genew; HGNC:7997; NRGL.
 CC MIM; 142445; --
 CC InterPro; IPR000561; EGF-like.
 CC Pfam; PF00008; EGF; 1.
 CC SMART; SM00181; EGF; 1.
 CC PROSITE; PS00022; EGF_1; 1.
 CC PROSITE; PS01186; EGF_2; FALSE_NEG.
 CC Growth factor; EGF-like domain; Transmembrane; Multigene family;
 CC Alternative splicing.
 CC TRANSMEM 76 100 INTERNAL SIGNAL SEQUENCE (POTENTIAL).
 CC DOMAIN 58 91 CYS-RICH.
 CC DOMAIN 211 232 SER/THR-RICH.
 CC DOMAIN 233 277 EGF-LIKE.
 CC DISULFID 237 251 BY SIMILARITY.
 CC DISULFID 245 265 BY SIMILARITY.
 CC DISULFID 267 276 BY SIMILARITY.
 CC SEQUENCE 296 AA; 31685 MW; 8041743217F/EB02 CRC64;
 SQ

Query Match

Best Local Similarity 7.9%; Score 273; DB 1; Length 296;

Pred. No. 1.2e-08;

Matches 62; Conservative 17; Mismatches 39; Indels 28; Gaps 4;
 QY 151 SGEYMKVSKLGNDSASANITIVESNIITGMPASTEG-----AYVSSSP----- 197
 DB 136 SSEAYTSPVSRQAQSES-EVQVTVGDKAVVFSFAAPTAKNRIFAFLPSTAPSPSP 194
 QY 198 -----IRISVSTEGANTSS-----STSTSTGTSHLVKCAEKEKTCVNGGECFMV 243
 DB 195 TRNPEVTRPKSATQPTETNLQAPKLSTSTGTSTSHLVKCAEKEKTCVNGGECFMV 254
 QY 244 KDLNSPNSRYLCKCPGFTGARGTENV 269
 DB 255 KDLNSPNSRYLCKCPNEFTGRCQNVV 280
 RESULT 11
 NP14_RAT
 ID NP14_RAT STANDARD; PRT; 704 AA.
 AC P4177;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Nucleolar phosphoprotein p130 (Nucleolar 130 kDa protein) (140 kDa
 DE nucleolar phosphoprotein) (Nopp140) (Nucleolar and coiled-body
 DE phosphoprotein 1).
 GN NOLC1.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 292-309 AND 563-601.
 RC TISSUE=Liver;
 RX MEDLINE=92323542; PubMed=1623516;
 RA Meier U.T., Blobel G.;
 RT "Nopp140 shuttles on tracks between nucleolus and cytoplasm.";
 RL Cell 70:127-138(1992).
 RN [2]
 RP INTERACTION WITH NOP5 AND FIBRILLARIN.
 RX MEDLINE=20143579; PubMed=10679015;
 RA Yang Y., Isaac C., Wang C., Dragon F., Pogacic V., Meier U.T.;
 RT "Conserved composition of mammalian box H/ACA and box C/D small
 RT nucleolar ribonucleoprotein particles and their interaction with the
 RT common factor Nopp140.";
 RL MOL. Biol. Cell 11:567-577(2000).
 CC -1- FUNCTION: RELATED TO NUCLEOGENESIS, MAY PLAY A ROLE IN THE
 CC MAINTENANCE OF THE FUNDAMENTAL STRUCTURE OF THE FIBRILLAR CENTER
 CC AND DENSE FIBRILLAR COMPONENT IN THE NUCLEOLUS. IT HAS INTRINSIC
 CC GTPASE AND ATPASE ACTIVITIES. MAY PLAY AN IMPORTANT ROLE IN
 CC TRANSCRIPTION CATALYZED BY RNA POLYMERASE I (BY SIMILARITY).
 CC -1- SUBUNIT: Interacts with Dkl1/Nap57, Nop5/Nap65 and fibrillarin.
 CC -1- SUBCELLULAR LOCATION: SHUTTLES ON CURVILINEAR TRACKS BETWEEN
 CC NUCLEOLUS AND CYTOPLASM. THESE TRACKS EXTEND FROM THE DENSE
 CC FIBRILLAR COMPONENT OF THE NUCLEOLUS ACROSS THE NUCLEOPLASM TO
 CC A LIMITED NUMBER OF NUCLEAR PORE COMPLEXES.
 CC -1- PTM: THIS PROTEIN UNDERGOES RAPID AND MASSIVE PHOSPHORYLATION AND
 CC DEPHOSPHORYLATION ON CK-II AND PKC SITES. Nopp140 IS ONE OF THE
 CC MOST PHOSPHORYLATED PROTEINS IN THE CELL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M94287; AAA41718.1; --
 CC EMBL; M94288; AAA41719.1; --
 CC Nucleolar protein; Phosphorylation; Repeat; GTP-binding; ATP-binding.
 CC DOMAIN 84 570 11 X 12 AA APPROXIMATE REPEATS OF AN
 CC ACIDIC SERINE CLUSTER.
 CC REPEAT 84 95 ACIDIC SERINE CLUSTER 1.
 FT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2003, 10:37:10 ; Search time 88 Seconds
(without alignments)
1566.426 Million cell updates/sec

Title: US-10-022-609-11

Perfect score: 3470
Sequence: 1 ARAPQRGRSLSPSRDKLFPN.....RLADSRNPAGRFSTQEEIQ 619

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rhodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2800.	80.7	782	11 Q9ESA5	Q9esa5 rattus norv
2	2248.5	64.8	695	11 Q9ESB0	Q9esb0 rattus norv
3	2141	61.7	700	11 Q9ESB1	Q9esb1 rattus norv
4	2026.5	58.4	461	11 Q35947	Q35947 mesocricetu
5	1282.5	37.0	298	11 Q9ESA9	Q9esa9 rattus norv
6	1024	29.5	241	6 Q07112	Q07112 bos taurus
7	935.5	27.0	317	11 Q9ESA3	Q9esa3 rattus norv
8	929	26.8	342	11 Q9ESA1	Q9esa1 rattus norv
9	921	26.5	323	11 Q9ESA2	Q9esa2 rattus norv
10	273	7.9	296	4 Q96IB3	Q96ib3 homo sapien
11	269.5	7.8	111	11 Q9ESA8	Q9esa8 rattus norv
12	269	7.8	256	11 Q9ESA6	Q9esa6 rattus norv
13	266	7.7	136	11 Q9ESA7	Q9esa7 rattus norv
14	192	5.5	1514	5 Q8SY55	Q8sy55 drosophila
15	189.5	5.4	1324	5 Q8T2N1	Q8tzn1 dictyosteli
16	188.5	5.4	1106	4 Q8WX93	Q8wx93 homo sapien

17	188.5	5.4	5476	5 Q9NJ17	Q9nj17 drosophila
18	188.5	5.4	5533	5 Q9VPL2	Q9vpl2 drosophila
19	188.5	5.4	5533	5 Q9UC63	Q9uc63 drosophila
20	188.5	5.4	5554	5 Q9NHN1	Q9nhn1 drosophila
21	188.5	5.4	5560	5 Q9VPL1	Q9vpl1 drosophila
22	187.5	5.4	1565	5 Q8TIM2	Q8tim2 dictyosteli
23	187	5.4	3257	5 Q9V736	Q9v736 drosophila
24	180.5	5.2	5327	5 Q76891	Q76891 drosophila
25	180	5.2	897	11 Q70495	Q70495 mus musculus
26	179.5	5.2	2254	10 Q9LN02	Q9ln02 arabidopsis
27	177	5.1	2006	5 Q9WFE2	Q9wfe2 drosophila
28	172	5.0	3507	5 Q23587	Q23587 caenorhabdi
29	171	4.9	3111	5 Q9VHI0	Q9vhi0 drosophila
30	170	4.9	990	13 Q91803	Q91803 xenopus lae
31	166.5	4.8	1172	5 Q9VE53	Q9ve53 drosophila
32	166.5	4.8	1262	4 Q9UO40	Q9u040 homo sapien
33	166.5	4.8	1459	5 Q17084	Q17084 caenorhabdi
34	166.5	4.8	1464	5 Q61802	Q61802 caenorhabdi
35	166.5	4.8	2296	4 Q9UHA8	Q9uha8 homo sapien
36	166.5	4.8	2752	4 Q9UQ35	Q9uq35 homo sapien
37	165.5	4.8	700	4 Q9BUV3	Q9buv3 homo sapien
38	165.5	4.8	1164	5 Q9VI61	Q9vi61 drosophila
39	165	4.8	1228	11 Q9JKB5	Q9jkb5 rattus norv
40	165	4.8	1529	5 Q9QC2	Q9qc2 dictyosteli
41	164.5	4.7	1031	13 Q90VM2	Q90vm2 brachydanio
42	164.5	4.7	1097	4 Q8TDW7	Q8tdw7 homo sapien
43	162.5	4.7	6632	5 Q17362	Q17362 caenorhabdi
44	162.5	4.7	6632	5 Q01761	Q01761 caenorhabdi
45	161	4.6	1492	5 Q8SSU1	Q8ssul dictyosteli

ALIGNMENTS

RESULT 1

ID	Q9ESA5	PRELIMINARY;	PRT;	782 AA.
AC	Q9ESA5;			
DT	01-MAR-2001 (Tremblrel. 16, Created)			
DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)			
DE	Glial growth factor beta 1a (Fragment).			
GN	NRG1.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=SPINAL CORD/BRAIN STEM;			
RA	Carroll S.L., Stonecypher M.S., Anderson K.D., Pearson R.J. Jr.,			
RA	Frohnert P.W.;			
RT	"Structural and Functional Diversity of Glial Growth Factor Isoforms			
RT	Expressed in Regenerating Peripheral Nerve and Associated Neurons.";			
RL	Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF194993; AAG28433.1; -			
DR	HSSP; Q12784; 1HRE.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR002114; HPr_Serp_site.			
DR	InterPro; IPR003599; Ig.			
DR	InterPro; IPR003598; Ig_c2.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR002154; Neuregulin.			
DR	Pfam; PF00008; EGF; 1.			
DR	Pfam; PF00047; Ig; 1.			
DR	Pfam; PF02158; Neuregulin; 1.			
DR	PRINTS; PR01089; NEUREGULIN.			
DR	SMART; SM00181; EGF; 1.			
DR	SMART; SM00001; EGF-like; 1.			
DR	SMART; SM00409; IG; 1.			
DR	SMART; SM00408; IGC2; 1.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.			

FT DOMAIN 178 222 EGF-LIKE.
 FT DISULFID 57 112 BY SIMILARITY.
 FT DISULFID 182 196 BY SIMILARITY.
 FT DISULFID 190 210 BY SIMILARITY.
 FT DISULFID 212 221 BY SIMILARITY.
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 461 AA; 50890 MW; 935C9560F7148336 CRC64;

Query Match 58.4%; Score 2026.5; DB 11; Length 461;
 Best Local Similarity 91.5%; Pred. No. 1.5e-131;
 Matches 388; Conservative 15; Mismatches 20; Indels 1; Gaps 1;

QY 45 MSERKEGRGKGGKKGKRGKSPALPRKMKQSOAAGSKVLRLCEYS 104
 DB 1 MSERKEGRGKGGKKGKDRGRGKPAEGENPSPALPRKMKIQSAAGSKVLRLCEYS 60
 105 SEYSSLRFWKNGNLRNKKPNQIKQKPKSELINKASLADSGEYMKVSKLGN 164
 61 SEYPELRFKFWKNGSELNKRTPQIKLQKPKSELINKASLADSGEYMKVSKLGN 120
 165 DSASANTIVSENEIITGMPASTEGAYVSESPRISVSTEGANTSSSTSTGTGSHLV 224
 121 DSASANTIVDSNEFITGMPASTERAYVSESPRISVSTEGANTSSSTSTGTGSHLV 180
 225 KCAEKETFCVNGGECFVWKDLSNPSRYLCKCPGFTGARTENVPKQNOEKAEELYQ 284
 181 KCAEKETFCVNGGECFVWKDLSNPSRYLCKCPGFTGARTENVPKQNOEKAEELYQ 240
 285 KRVLTGTGICALLVVGIMCVAYCKTKKQKRLHRLRQSLRSERNMMNTANGPHHPN 344
 241 KRVLTGTGICALLVVGIMCVAYCKTKKQKRLHRLRQSLRSERNMMNTANGPHHPN 300
 345 PPEPVOLVNOYVKNVISSEHIVERAEETSESTSHVSTAHHSTVTPSPHSWNGHT 404
 301 PPEPVOLVNOYVKNVISSEHIVERAEETSESTSHVSTAHHSTVTPSPHSWNGHT 360
 405 ESILSESHSVIMSSVSNRSHSSPTGPRGLNGTGGPRECSFLRHARETDPDSYRDSPH 464
 361 ESIVSESHSVIMSSVSNRSHSSPTGPRGLNGTGGPRECSFLRHARETDPDSYRDSPH 419
 465 SERV 468
 420 SERH 423

RESULT 5

ID Q9ESA9 PRELIMINARY; PRT; 298 AA.

AC Q9ESA9;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE SHDF neuregulin alpha 2b (Fragment).
 GN NG1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BDIX;
 RA Carroll S.L., Anderson K.D., Frohnert P.W.;
 RT "Structural and Functional Diversity of SMDF Neuregulin Splice
 Variants Expressed in the Adult Rat Nervous System."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF194440; AAG28429.1; -;
 DR HSSP; Q12784; 1HRE.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002114; HPr_SerP_site.
 DR InterPro; IPR002154; Neuregulin.

DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF02158; Neuregulin; 1.
 DR PRINTS; PRO1089; NEUREGULIN.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00589; PTS_HPR_SER; UNKNOWN_1.
 FT NON_TER 1
 FT NON_TER 298
 SQ SEQUENCE 298 AA; 32851 MW; BD76F014C2B33026 CRC64;

Query Match 37.0%; Score 1282.5; DB 11; Length 298;
 Best Local Similarity 93.0%; Pred. No. 1.4e-80;
 Matches 239; Conservative 12; Mismatches 5; Indels 1; Gaps 1;

QY 212 STSTSTGTSHLVKCAEKEKTCVNGGECFVWKDLSNPSRYLCKCPGFTGARTENVP 271
 DB 38 STSTSTGTSHLVKCAEKEKTCVNGGECFVWKDLSNPSRYLCKCPGFTGARTENVP 97
 272 KYQNOEKAEELYQKRVLTGTGICALLVVGIMCVAYCKTKKQKRLHRLRQSLRSERN 331
 98 KVTQEKAEELYQKRVLTGTGICALLVVGIMCVAYCKTKKQKRLHRLRQSLRSERN 157
 332 NMNTANGPHHPNPPENPVOLVNOYVKNVISSEHIVERAEETSESTSHVSTAHHSTV 391
 158 NLVNTANGPHHPNPPENPVOLVNOYVKNVISSEHIVERAEETSESTSHVSTAHHSTV 217
 392 TQTPSHSWNGHTSESTSHVSTAHHSTV 451
 218 TQTPSHSWNGHTSESTSHVSTAHHSTV 276
 452 ARETDPDSYRDSPHSERV 468
 277 ARETDPDSYRDSPHSERH 293

RESULT 6
 Q07112
 ID Q07112 PRELIMINARY; PRT; 241 AA.
 AC Q07112;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE Neuregulin-1, glial growth factor 5 isoform precursor (GGFBPP5).
 GN NRGI OR GGF.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE-POSTERIOR PITUITARY.
 RX MEDLINE-93205115; PubMed-8096067;
 RA Marchionni M.A., Goodearl A.D.J., Chen M.S., Birmingham-McDonogh O.,
 RA Kirk C., Hendricks M., Danehy F., Misumi D., Sudhalter J.,
 RA Kobayashi K., Wroblewski D., Lynch C., Baldassarre M., Hiles I.,
 RA Davis J.B., Hsuan J.J., Totty N.F., Otsu M., McBurney R.N.,
 RA Waterfield M.D., Stroobant P., Gwynne D.
 RT "Glial growth factors are alternatively spliced erbB2 ligands
 expressed in the nervous system."
 RL Nature 362:312-318(1993).
 CC -1- FUNCTION: DIRECT LIGAND FOR ERBB3 AND ERBB4 TYROSINE KINASE
 RECEPTORS. CONCOMITANTLY RECRUITS ERBB1 AND ERBB2 CORECEPTORS,
 CC RESULTING IN LIGAND-STIMULATED TYROSINE PHOSPHORYLATION AND
 CC ACTIVATION OF THE ERBB RECEPTORS (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED (PROBABLE).
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN NERVOUS SYSTEM; SPINAL CORD AND
 CC BRAIN.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST FIVE DIFFERENT ISOFORMS; GGFBPPI,
 CC GGFBP2, GGFBP3, GGFBP4 AND GGFBP5 (SHOWN HERE); ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.


```

DR SMART; SM00409; IG; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00410; IG_1like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
KW Immunoglobulin domain.
FT NON_TER
SQ SEQUENCE 323 AA; 35358 MW; C7DF153A939A80C8 CRC64;

      26.5%; Score 921; DB 11; Length 323;
Query Match 65.6%; Pred. No. 1.2E-55;
Matches 198; Conservative 12; Mismatches 38; Indels 54; Gaps 5

Qy 12 PSRDKLPNPINRALGPSNPAPRAVVR-----SVSGEMSERKEGRK-----GK 56
Db 16 PSTDQ-----PGDPAPLYKLVHQVAVKAGLKKDLSLTAVLDTWGHPAPPSGR 65
Qy 57 GKKEK-----GSGKKPEAAGSQSP-----ALPPLRKEMK 87
Db 66 LKEDSRVIFFMEDPDANSSGRAPPAPRASFPPLGTGRNLKEVSRVLCRCALPPLRKEM 125
Qy 88 SQESAAGKLVLRCTSEYSYSLRFKWKNGNELNRKNKPNIKIQKKPGKSELIRNKAS 147
Db 126 SQESAAGKLVLRCTSEYSYSLRFKWKNGNELNRKNKPNIKIQKKPGKSELIRNKAS 185
Qy 148 LAQSGYMKVVISKLGNDASANITTVESNEIITGMPASTGAYVSESPIRISVSTEGA 207
Db 186 PADSGEYMKVVISKLGNDASANITTVESNEFITGMPASTATRVSSSPIRISVSTEGA 245
Qy 208 NTSSTSTSTGTSHLVKCAEKETFCVNGGCEFMWRDLSNPISYLCKCQPGFTGARCTE 267
Db 246 NTSSTSTSTGTSHLVKCAEKETFCVNGGCEFTVRDLSNPISYLCKCQPFETGDRCON 305

```

```

RESULT 10
Q96IB3      PRELIMINARY;          PRT;    296 AA.
ID Q96IB3;
AC AC Q96IB3;
DT DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE DE Neuregulin 1.
OS OS Homo sapiens (Human).
OC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;
RX RX [1]
RN RN SEQUENCE FROM N.A.
RP RP TISSUE-BRAIN;
RC RA Strausberg R.;
RL RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR DR EMBL; BC007675; AAH07675.1; -.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF_1
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00032; EGF_1; UNKNOWN_1.
SQ SQ SEQUENCE 296 AA; 31699 MW; 0F5C48C8465B6649 CRC64;

Query Match              7.98;   Score 273; DB 4; Length 296;
Best Local Similarity    42.5%; Pred. No. 4.4e-11;
Matches 62; Conservative 17; Mismatches 39; Indels 28; Gaps 4;

Qy 151 SGEYMKVISKLGNDASANITIVESNEIITGMPASTE-----AYVSSESP----- 197
Db 136 SSEAYTSPVSRQSES-EVQTVGQRNAVSVFEPSAAPTKNRIFAESFLPSTAPSFSP 194

Qy 198 -----IRISVTREGANTSS-----STSTSTGTSHLVKCAEKTEFCVNGECFMV 243
Db 195 TRNPEVRTPKSAPOPTETNLTOPAKLTSTSTGTSHLVKCAEKTKTCVNGECCFMV 254
```

```

QY 244 KDLNPSRYLCKCPGFTGARTENV 269
DB 255 KDLNPSRYLCKCPNEFTGDRCONV 280

RESULT 11
Q9ESA8 PRELIMINARY; PRT; 111 AA.
AC Q9ESA8;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SMDF neuregulin beta 2 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194441; AAG28430.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000886; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 111
SQ SEQUENCE 111 AA; 12198 MW; C88BB70584C9F8C CRC64;

Query Match 7.8%; Score 269.5; DB 11; Length 111;
Best Local Similarity 74.6%; Pred. No. 2.2e-11;
Matches 53; Conservative 4; Mismatches 11; Indels 3; Gaps 1;

QY 212 STSTTTGTHLVKCAEKFTCVNGGCFVVKDLNPSRYLCKCPGFTGARTENVPM 271
DB 44 STTSTTSHLHKCAEKFTCVNGGCFVVKDLNPSRYLCKCPNEFTGDRCONVY-- 101

QY 272 KVNQOEKAEEL 282
DB 102 -WASYKAEEL 111

RESULT 12
Q9ESA6 PRELIMINARY; PRT; 256 AA.
AC Q9ESA6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SMDF neuregulin beta 3 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF194443; AAG28432.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 256 AA; 27335 MW; C8C08DEA68D33E39 CRC64;

Query Match 7.8%; Score 269; DB 11; Length 256;
Best Local Similarity 41.8%; Pred. No. 6.9e-11;
Matches 61; Conservative 11; Mismatches 46; Indels 28; Gaps 3;

QY 151 SGEYMKYVISKLGNDASANITIVESNEIITGMPA-----STEGAYVSS 194
DB 96 SSEAYTSPVSR-AQSEAGAHVTVQGDHAAVAASEPSAVPTRKNRLSAFPFPHSTAPPFPSP 154

QY 195 ESPRISVSTEGANTSS-----STSTTTGTHLVKCAEKFTCVNGGCFV 243
DB 155 ARTPEVTRPKSGTQPTTETNLQAPKLSTSTTGTSHLHKCAEKFTCVNGGCFV 214

QY 244 KDLNPSRYLCKCPGFTGARTENV 269
DB 215 KDLNPSRYLCKCPNEFTGDRCONV 240

RESULT 13
Q9ESA7 PRELIMINARY; PRT; 136 AA.
AC Q9ESA7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SMDF neuregulin beta 4 (Fragment).
GN NRGI.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RC TISSUE-AXOTOMIZED LUMBAR DORSAL ROOT GANGLION/SPINAL CORD;
RA Carroll S.L., Anderson K.D., Frohnert P.W.;
RT "Structural and Functional Diversity of SMDF Neuregulin Splice
RT Variants Expressed in the Adult Rat Nervous System.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF194442; AAG28431.1; -.
DR HSSP; Q12784; 1HRE.
DR InterPro; IPR000561; EGF-like.
DR Pfam; PF00008; EGF; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00001; EGF-like; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 136
SQ SEQUENCE 136 AA; 15235 MW; 116CBC91D6E8AEF5 CRC64;

Query Match 7.7%; Score 266; DB 11; Length 136;
Best Local Similarity 69.6%; Pred. No. 4.9e-11;
Matches 55; Conservative 2; Mismatches 12; Indels 10; Gaps 2;

QY 212 STSTTTGTHLVKCAEKFTCVNGGCFVVKDLNPSRYLCKCPGFTGARTENVPM 271
DB 44 STTSTTSHLHKCAEKFTCVNGGCFVVKDLNPSRYLCKCPNEFTGDRCONV 102

QY 272 -----KVNQOEKAE 281
DB 103 ASFYMTSRKKQETEKPLE 121

RESULT 14

```

[illegible]